

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:22:43 ; Search time 3867 Seconds
(without alignments)
10799.722 Million cell updates/sec

Title: US-09-763-748-1
Perfect score: 1435
Sequence: 1 ctcgscgscgtggaacca.....gataataagatatacagc 1435

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_hcg_hum:*

31: em_hcg_inv:*

32: em_hcg_other:*

33: em_hcg_mus:*

34: em_hcg_pln:*

35: em_hcg_rtd:*

36: em_hcg_man:*

37: em_hcg_vtl:*

38: em_sy:*

39: em_hcgo_hum:*

40: em_hcgo_mus:*

41: em_hcgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1435	100.0	1435	6	AR098746
2	1435	100.0	1435	6	HUMTRADD
3	1435	100.0	1441	6	I27061
4	1433.4	99.9	1475	9	BC004491
5	1243.8	86.7	1833	9	AK090673
6	1005.8	70.1	152953	9	AC074143
7	1005.8	70.1	209574	9	AC040160
8	990.8	69.0	228516	2	AC008594
9	512.2	35.7	579	9	HS011616
10	434.4	30.3	208936	2	AC010821
11	427.6	29.8	1384	6	I27062
12	398.4	27.8	927	9	HS011615
13	282.6	19.7	237933	2	AC124713
14	272	19.0	180835	2	AC120484
15	200	13.9	200	6	AR211525
16	175.8	12.3	334	10	AF517017
17	175.4	12.2	740	9	HS033101
18	175.4	12.2	743	9	HS033648
19	172.2	12.0	1184	9	HS033449
20	165.8	11.6	731	9	HS032463
21	164.6	11.5	704	9	HS031614
22	127	8.9	552	10	BC028542
23	111.4	7.8	882	5	AF231014
24	103.6	7.2	180835	2	AC120484
25	77.4	5.4	155662	2	AC112388
26	72.8	5.1	3060	3	AF027735
27	72	5.0	123020	9	AF429315
28	70.6	4.9	125020	9	AF429315
29	69.8	4.9	1279	11	PM2H12G
30	69.4	4.8	40476	1	SC075A
31	69.2	4.8	2793	6	AR009990
32	69.2	4.8	2793	6	I95876
33	65.2	4.5	34023	1	SC262
34	63.8	4.4	1150	14	HS04ULR3
35	63.8	4.4	1926	6	AX107940
36	63.8	4.4	2580	6	AR108994
37	63.8	4.4	5452	6	AR083151
38	63.8	4.4	5452	12	U02454
39	63.8	4.4	9600	6	A92665
40	63.8	4.4	10596	6	AR158345
41	63.8	4.4	10596	6	I25041
42	63.8	4.4	10596	6	I30503
43	63.8	4.4	10737	12	XXU02428
44	63.8	4.4	10850	12	U02455
45	63.8	4.4	172281	14	EBV

ALIGNMENTS

RESULT 1

AR098746 1435 bp DNA

LOCUS AR098746

DEFINITION Sequence 1 from patent US 6077672.

ACCESSION AR098746

VERSION AR098746.1 GI:12808512

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1435)

AUTHORS Monté, B.P. and Cowser, L.M.

TITLE Antisense modulation of TRAD expression

JOURNAL Patent: US 6077672-A 1 20-JUN-2000;

FEATURES Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a

source 1..1435
/organism="unknown"
BASE COUNT 245 a 437 c 494 g 259 t
ORIGIN

Query Match 100.0%; Score 1435; DB 6; Length 1435;
Best Local Similarity 100.0%; Pred. No. 2.3e-220;
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGGGGGGTGGGAACCCAGGCCCCCGGAGGCGGCGAGAGTGAATGGAGCTGGG 60
DB 1 CTGGGGGGGGTGGGAACCCAGGCCCCCGGAGGCGGCGAGAGTGAATGGAGCTGGG 60
QY 61 CAAAATGGGCGAAGAGTGGTGGGAGCGCATACCTGTTTGGAGTCTTGCTGAGAC 120
DB 61 CAAAATGGGCGAAGAGTGGTGGGAGCGCATACCTGTTTGGAGTCTTGCTGAGAC 120
QY 121 AAGGTGGTCTGTGGATGCTTACGCGCACCCCGACAGAAAGTGGCAGTGTACAGGCT 180
DB 121 AAGGTGGTCTGTGGATGCTTACGCGCACCCCGACAGAAAGTGGCAGTGTACAGGCT 180
QY 181 CTGCAAGGCTGCTTGGCAGAGAGCGGCGGAGCGCGGAGCGTGTGAGATGCTGAAGATC 240
DB 181 CTGCAAGGCTGCTTGGCAGAGAGCGGCGGAGCGCGGAGCGTGTGAGATGCTGAAGATC 240
QY 241 CACCGGAGGAGCCCGGAGCTGATGTCAGACTGCGATTCTGGGGGCGGACCTGTGGC 300
DB 241 CACCGGAGGAGCCCGGAGCTGATGTCAGACTGCGATTCTGGGGGCGGACCTGTGGC 300
QY 301 GCGCTTCTCGCGGCTTACCGCGAGAGGGGCGCTGCGCGCGCGCTGAGAGAGCCTGGCG 360
DB 301 GCGCTTCTCGCGGCTTACCGCGAGAGGGGCGCTGCGCGCGCGCTGAGAGAGCCTGGCG 360
QY 361 GCGGCGCTGCGGCGAGACTCGGTGCGGCTGCACTGGAAGTGGCGGCGGCGGAGCGG 420
DB 361 GCGGCGCTGCGGCGAGACTCGGTGCGGCTGCACTGGAAGTGGCGGCGGCGGAGCGG 420
QY 421 CTGGAAGCTTGTGCGGAGAGGAGCGCTGTTGATTGATCTTACGCCAGCGCC 480
DB 421 CTGGAAGCTTGTGCGGAGAGGAGCGCTGTTGATTGATCTTACGCCAGCGCC 480
QY 481 GACCGGCTCGCGGAGTGAAGACTGCTGAGCTGAGAGATGCGCTGGAATCTGAAGTGC 540
DB 481 GACCGGCTCGCGGAGTGAAGACTGCTGAGCTGAGAGATGCGCTGGAATCTGAAGTGC 540
QY 541 GCGTGGGGGGGCGGCGGAGCGGAGGAGTGGCTTGGCGCCCTTGAAGCCCCCGGTG 600
DB 541 GCGTGGGGGGGCGGCGGAGCGGAGGAGTGGCTTGGCGCCCTTGAAGCCCCCGGTG 600
QY 601 CCTCTCTGTGCGAGGAGTGAAGCGCGCGCGCGCGCACCTGCGCAGATTTTCTGTTT 660
DB 601 CCTCTCTGTGCGAGGAGTGAAGCGCGCGCGCGCGCACCTGCGCAGATTTTCTGTTT 660
QY 661 CAGGGTCAAGCTGTAGTGAATCGGCGGCTGAGCTGAAGAGCCAGACGTTTGGCGCGC 720
DB 661 CAGGGTCAAGCTGTAGTGAATCGGCGGCTGAGCTGAAGAGCCAGACGTTTGGCGCGC 720
QY 721 TCTGTGGGCTCAAAATGGGCGAAGTGGGCGCTCACTGAGGAGGCTGCGGGGCGCTG 780
DB 721 TCTGTGGGCTCAAAATGGGCGAAGTGGGCGCTCACTGAGGAGGCTGCGGGGCGCTG 780
QY 781 CCGGAGCCCGGCGCTGAGCTGCGGCTGAGAGTGAAGAGGCGGAGGCTGTACAGACAG 840
DB 781 CCGGAGCCCGGCGCTGAGCTGCGGCTGAGAGTGAAGAGGCGGAGGCTGTACAGACAG 840
QY 841 GCGTTCACAGCTGTGCGGCGCTTGTGAGAGCGCGAGAGCGCGCGCACTGCTGACGCG 900
DB 841 GCGTTCACAGCTGTGCGGCGCTTGTGAGAGCGCGAGAGCGCGCGCACTGCTGACGCG 900
QY 901 CTGGTGGAGGCACTGAGAGGAAAGGCTCACAGGCTTGGCAGAGGACTTGTGCGGCGCTG 960
DB 901 CTGGTGGAGGCACTGAGAGGAAAGGCTCACAGGCTTGGCAGAGGACTTGTGCGGCGCTG 960

QY 961 ACCGATCCCAATGAGCGGCTGCGCTAGACAGAGGGTGGAGCGGCTTTGGAGAACTTGG 1020
DB 961 ACCGATCCCAATGAGCGGCTGCGCTAGACAGAGGGTGGAGCGGCTTTGGAGAACTTGG 1020
QY 1021 ATGGCTTAAAGGTTCTTCTGCGGCTATTGCTGAACCCCTGTTCATTCACGGGACCTTGA 1080
DB 1021 ATGGCTTAAAGGTTCTTCTGCGGCTATTGCTGAACCCCTGTTCATTCACGGGACCTTGA 1080
QY 1081 AACTCCACTTGGCCATATGCTGAGACTGAGCGGCGGAGAGTGAATGCTTCCCGAGAG 1140
DB 1081 AACTCCACTTGGCCATATGCTGAGACTGAGCGGCGGAGAGTGAATGCTTCCCGAGAG 1140
QY 1141 CCAGACCACTGGGGGTGATCATTTGGGAGTTCTGCTCAGGATCTTTGATAGAGTGGG 1200
DB 1141 CCAGACCACTGGGGGTGATCATTTGGGAGTTCTGCTCAGGATCTTTGATAGAGTGGG 1200
QY 1201 GTGGGGGGGACTTGTCTTGGAGATCAGCTTCACTTCCATCCAGAGCGGGGCTTGA 1260
DB 1201 GTGGGGGGGACTTGTCTTGGAGATCAGCTTCACTTCCATCCAGAGCGGGGCTTGA 1260
QY 1261 CAGCGAGCCCTTACAGTTTCACTCACTGAAGCACTTGAATCTTGGTGTCTGGACTTGA 1320
DB 1261 CAGCGAGCCCTTACAGTTTCACTCACTGAAGCACTTGAATCTTGGTGTCTGGACTTGA 1320
QY 1321 CTGGGTGCTGAGATCTGCACTGAGTGAATGAACAGAGATCAATCTTGCCTGCCACG 1380
DB 1321 CTGGGTGCTGAGATCTGCACTGAGTGAATGAACAGAGATCAATCTTGCCTGCCACG 1380
QY 1381 TCACACTAGCGGTGGAGCCCGGATGTTAAGCATGATTAATGAATTAACGCG 1435
DB 1381 TCACACTAGCGGTGGAGCCCGGATGTTAAGCATGATTAATGAATTAACGCG 1435

RESULT 2

HUMTRADD 1435 bp mRNA linear PRI 26-APR-1996
LOCUS Homo sapiens TNF receptor-1 associated protein (TRADD) mRNA, 3' end

DEFINITION

of cds.

ACCESSION

U41690.1 GI:808914

VERSION

L41690

KEYWORDS

TNF receptor; TRADD gene; tumor necrosis factor receptor; tumor necrosis factor receptor 1; tumor necrosis factor receptor-1 associated protein.

SOURCE

Homo sapiens.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.

REFERENCE

1 (bases 1 to 1435)

AUTHORS

Hsu, H., Xiong, J., and Goeddel, D.V.

TITLE

The TNF receptor 1-associated protein TRADD signals cell death and NF-kappa B activation

JOURNAL

Cell 81 (4), 495-504 (1995)

MEDLINE

95277836

PUBMED

7758105

FEATURES

source

1..1435

/organism="Homo sapiens"

/db_xref="taxon:9606"

1..1435

/gene="TRADD"

<1..987

/standard_name="TNF receptor-1 associated protein"

/codon_start=1

/product="tumor necrosis factor receptor type 1 associated protein"

/protein_id="AA98482.1"

/db_xref="GI:808915"

/translation="LAVGVQAPRRPGRGEMAGONGHEWVGSAYLVFVSSLDKVVLLSDAYHQKAVYVTRALQALALABSGSPDVLQMTKTRSPQULVQLRFGRCRGRF LRAIRGALRALDRSLAALADHSVPLDLRLRGAERLALADBERCSLTAQDP DRLEDELALIEDLRLKCGSGRGGDGVASPLQPPVPSISEVPPPPPAQTF LFOGQPVNRPPLSLKDOOTFARSVGLKMRKVSLSLQRCBALRDPALDSLAYEYREG

BASE COUNT 245 a 437 c 494 g 259 t
 LYPQAFQLLRRFVQAEGRATLQRLVEALEENELTSLAEDLLGLDPMNGIA"
 ORIGIN

Query Match 100.0%; Score 1435; DB 9; Length 1435;
 Best Local Similarity 100.0%; Pred. No. 2,3e-220;
 Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGGGCGTGGGAACCCAGAGCCCGGAGGGGCGGAGAGGTGAGATGGAGCTGG 60
 DB 1 CTGGGGGCGTGGGAACCCAGAGCCCGGAGGGGCGGAGAGGTGAGATGGAGCTGG 60
 QY 61 CAAATGGGACGAAAGTGGGTGGGCGAGCCATACCTGTTTGGAGTCTCTCGTGGAC 120
 DB 61 CAAATGGGACGAAAGTGGGTGGGCGAGCCATACCTGTTTGGAGTCTCTCGTGGAC 120
 QY 121 AAGGTGCTCTGTGGATGCTTACGCGCACCCCGACGAAAGTGGAGTGTACAGGCT 180
 DB 121 AAGGTGCTCTGTGGATGCTTACGCGCACCCCGACGAAAGTGGAGTGTACAGGCT 180
 QY 181 CTGAGAGCTGCTTGGGACGAGAGGGGGGAGCCCGGACGCTGCGAGATGCTGAAGATC 240
 DB 181 CTGAGAGCTGCTTGGGACGAGAGGGGGGAGCCCGGACGCTGCGAGATGCTGAAGATC 240
 QY 241 CACCGCAGCAGCCGACAGCTGATGCTGAGCTGCGATTCTGCGGGCGGACCCCTGTGGC 300
 DB 241 CACCGCAGCAGCCGACAGCTGATGCTGAGCTGCGATTCTGCGGGCGGACCCCTGTGGC 300
 QY 301 CGCTTCTCGCGCGCTTACCGCGAGGGGGCGCTGCGCGCGCTGCGAGAGAGCCTGGCG 360
 DB 301 CGCTTCTCGCGCGCTTACCGCGAGGGGGCGCTGCGCGCGCTGCGAGAGAGCCTGGCG 360
 QY 361 GCGCGGCTCGCGCGGACGACTGCTGCTGCGGCTGCAATGAGAGCTGCGGGCGGAGAGG 420
 DB 361 GCGCGGCTCGCGCGGACGACTGCTGCTGCGGCTGCAATGAGAGCTGCGGGCGGAGAGG 420
 QY 421 CTGAGAGCTTGTCTGGGAGAGAGAGCGCTGTTTGAATTCATCTTACCCAGAGAGCC 480
 DB 421 CTGAGAGCTTGTCTGGGAGAGAGAGCGCTGTTTGAATTCATCTTACCCAGAGAGCC 480
 QY 481 GACCGGCTCGCGGATGAAGAACTGGCTGAGCTGAGAGATGCTGCGAAATCTGAAGTGC 540
 DB 481 GACCGGCTCGCGGATGAAGAACTGGCTGAGCTGAGAGATGCTGCGAAATCTGAAGTGC 540
 QY 541 GGGCTCGGGGGGCGGGGGTGGAGAGGGAGGTGCTTGGGCGCCCTTTCAGACCCCGGTG 600
 DB 541 GGGCTCGGGGGGCGGGGGTGGAGAGGGAGGTGCTTGGGCGCCCTTTCAGACCCCGGTG 600
 QY 601 CCTCTCTGTGAGAGGTGAAGCGCGCGCGCGGACCTGCGCAGACTTTTCTGTTTC 660
 DB 601 CCTCTCTGTGAGAGGTGAAGCGCGCGCGCGGACCTGCGCAGACTTTTCTGTTTC 660
 QY 661 CAGGGTCAAGCTGTAGTGAATCGGCGGCTGAGCTTGAAGAACCAACAGAGTTCGCGGC 720
 DB 661 CAGGGTCAAGCTGTAGTGAATCGGCGGCTGAGCTTGAAGAACCAACAGAGTTCGCGGC 720
 QY 721 TCTGTGGGTCTCAAAATGGCGGAAGGTGGGGGCTTACTGACGAGAGGTGCGGGCGCTG 780
 DB 721 TCTGTGGGTCTCAAAATGGCGGAAGGTGGGGGCTTACTGACGAGAGGTGCGGGCGCTG 780
 QY 781 CCGGACCCGCGGCTGAGACTCGCTGGCTACAGATACAGAGCGAGAGGACTGTACAGAG 840
 DB 781 CCGGACCCGCGGCTGAGACTCGCTGGCTTACAGATACAGAGCGAGAGGACTGTACAGAG 840
 QY 841 GCTTTCAGCTGCTGCGGCGCTTGTGAGAGCCGAGAGGCGCGCGGACGCTGACGCGC 900
 DB 841 GCTTTCAGCTGCTGCGGCGCTTGTGAGAGCCGAGAGGCGCGCGGACGCTGACGCGC 900
 QY 901 CTGTGTGAGGAGCACTGAGAGAGAGAGAGCTTACAGAGCTGCGAGAGAGATTTGTGGGCTG 960
 DB 901 CTGTGTGAGGAGCACTGAGAGAGAGAGAGCTTACAGAGCTGCGAGAGAGATTTGTGGGCTG 960
 QY 961 ACCGATCCCAATGGCGGCTGGCTGAGACAGGGGTGACGACGCTTTGAGAACTGG 1020

DB 961 ACCGATCCCAATGGCGGCTGGCTGAGACAGGGGTGACAGCAGCTTTGAGAACTGG 1020
 QY 1021 ATGGCTTAAAGGTTCCTTTCGCGGCTATTGCTGAACCCCTGTCCATTCAGAGGACCTGA 1080
 DB 1021 ATGGCTTAAAGGTTCCTTTCGCGGCTATTGCTGAACCCCTGTCCATTCAGAGGACCTGA 1080
 QY 1081 AACTCCATTGGGCTTATCTGTGGACCTGTGCGGGGCGAGAGTGTATTCCTTCCAGAG 1140
 DB 1081 AACTCCATTGGGCTTATCTGTGGACCTGTGCGGGGCGAGAGTGTATTCCTTCCAGAG 1140
 QY 1141 CCAAGACACTGGGGGTGCATCATTTGGGATTTGCTCAGAGTACTTTGATAGAGTGGG 1200
 DB 1141 CCAAGACACTGGGGGTGCATCATTTGGGATTTGCTCAGAGTACTTTGATAGAGTGGG 1200
 QY 1201 GTGGGGGGGACTTGTCTTGGAGATCAGCTTACCTTCCCATCCCGAAGCGGGGCTTA 1260
 DB 1201 GTGGGGGGGACTTGTCTTGGAGATCAGCTTACCTTCCCATCCCGAAGCGGGGCTTA 1260
 QY 1261 CAGCAGACCCCTTACAGTTTCACTCATGAGACACCTTATCTTGTGTCTTGGACCTTCA 1320
 DB 1261 CAGCAGACCCCTTACAGTTTCACTCATGAGACACCTTATCTTGTGTCTTGGACCTTCA 1320
 QY 1321 CTTGGGTCTGCAAGATCTGAGTGAAGTAAACAGGAATCAATCTTGCCTGCCCGAC 1380
 DB 1321 CTTGGGTCTGCAAGATCTGAGTGAAGTAAACAGGAATCAATCTTGCCTGCCCGAC 1380
 QY 1381 TCACACTCAGCGTGGGAGCCCGGAATGTTAAGCAATGATTAATTAAGTAAACAGG 1435
 DB 1381 TCACACTCAGCGTGGGAGCCCGGAATGTTAAGCAATGATTAATTAAGTAAACAGG 1435

RESULT 3

LOCUS 127061 127061 1441 bp DNA linear PAT 06-FEB-1997
 DEFINITION Sequence 1 from patent US 5563039.
 ACCESSION 127061
 VERSION 127061.1 GI:1817837
 KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1441)

AUTHORS Goeddel, D.V. and Hsu, H.

TITLE TNF receptor-associated intracellular signaling proteins and

JOURNAL Patent: US 5563039-A 1 08-OCT-1996;

FEATURES 1..1441

source Location/Qualifiers

BASE COUNT 251 a 437 c 494 g 259 t
 ORIGIN /organism="unknown"

Query Match 100.0%; Score 1435; DB 6; Length 1441;
 Best Local Similarity 100.0%; Pred. No. 2,3e-220;
 Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCGGCGCTGGGAACCCAGAGCCCGGAGGCGGCGGAGAGGTGAGATGGCAGCTGG 60
 DB 1 CTGGCGGCGCTGGGAACCCAGAGCCCGGAGGCGGCGGAGAGGTGAGATGGCAGCTGG 60
 QY 61 CAAATGGGACGAAAGTGGGTGGGCGAGCCATACCTGTTTGGAGTCTCTCGTGGAC 120
 DB 61 CAAATGGGACGAAAGTGGGTGGGCGAGCATACCTGTTTGGAGTCTCTCGTGGAC 120
 QY 121 AAGGTGCTCTGTGGATGCTTACGCGCACCCCGACGAAAGTGGAGTGTACAGGCT 180
 DB 121 AAGGTGCTCTGTGGATGCTTACGCGCACCCCGACGAAAGTGGAGTGTACAGGCT 180
 QY 181 CTGAGAGCTGCTTGGGACGAGAGGGCGGAGAGCCCGGACGCTGCTGAGATGCTGAAGATC 240
 DB 181 CTGAGAGCTGCTTGGGACGAGAGGGCGGAGAGCCCGGACGCTGCTGAGATGCTGAAGATC 240

```

QY 241 CACCGGAGGACCCGAGCTGATCTGTCAGACTCGATTCTGCGGGCGAGACCCCTGTGCG 300
    |||
Db 241 CACCGGAGGACCCGAGCTGATCTGTCAGACTCGATTCTGCGGGCGAGACCCCTGTGCG 300
QY 301 CGCTTCTCCGCGCCCTTACCGCGAGGGGGCGCTGCGCGCGCGCTGCAAGAGAGCTTGGCG 360
    |||
Db 301 CGCTTCTCCGCGCCCTTACCGCGAGGGGGCGCTGCGCGCGCGCTGCAAGAGAGCTTGGCG 360
QY 361 GCGCGGCTGCGCGGAGCTCGGTGCGCGCTGCAAGAGAGCTGCGCGCGCGCGCGCGAGCGG 420
    |||
Db 361 GCGCGGCTGCGCGGAGCTCGGTGCGCGCTGCAAGAGAGCTGCGCGCGCGCGCGCGAGCGG 420
QY 421 CTGGAAGCTTTGCTGCGGAGCGAGAGCGCGCTTTTGAAGTTGCACTTCCAGCCAGAGCCG 480
    |||
Db 421 CTGGAAGCTTTGCTGCGGAGCGAGAGCGCGCTTTTGAAGTTGCACTTCCAGCCAGAGCCG 480
QY 481 GACCGGCTCCGGGATGAAAGACTGCTGAGCTGAGAGATGCGTGGGAAATCTGAAGTGC 540
    |||
Db 481 GACCGGCTCCGGGATGAAAGACTGCTGAGCTGAGAGATGCGTGGGAAATCTGAAGTGC 540
QY 541 GCGTGGGGGGCCCGGGGTGGCGAGCGGGAGAGTGGCTTTCGGGCGCCCTTGCAGCGCCGGGTG 600
    |||
Db 541 GCGTGGGGGGCCCGGGGTGGCGAGCGGGAGAGTGGCTTTCGGGCGCCCTTGCAGCGCCGGGTG 600
QY 601 CCCTCTCTGTGCGAGGTGAAGCCCGCGCGCGCGCGCACTGCGCGAGCTTTTCTGTTTC 660
    |||
Db 601 CCCTCTCTGTGCGAGGTGAAGCCCGCGCGCGCGCGCACTGCGCGAGCTTTTCTGTTTC 660
QY 661 CAGGCTCAGCTTGTAGTGAATCGGCGCGCTGAGCTTGAAGAGCAAGAGCTTTCGCGCGC 720
    |||
Db 661 CAGGCTCAGCTTGTAGTGAATCGGCGCGCTGAGCTTGAAGAGCAAGAGCTTTCGCGCGC 720
QY 721 TCTGTGGGTCTCAATAGGCGGCAAGGTGGGGGCTCACTGAGGAGGCTGCGGGGGCTG 780
    |||
Db 721 TCTGTGGGTCTCAATAGGCGGCAAGGTGGGGGCTCACTGAGGAGGCTGCGGGGGCTG 780
QY 781 CCGGAGCCCGCGCTGAGCTGCTGCGCTTACGAGTACGAGCGCGAGGAGCTTGTACGAGAG 840
    |||
Db 781 CCGGAGCCCGCGCTGAGCTGCTGCGCTTACGAGTACGAGCGCGAGGAGCTTGTACGAGAG 840
QY 841 GCGTTCAGCTGCTGCGCGCGCTTCTGTCAGAGCGCGAGGGCGCGCGCGCACGCTGCAAGCG 900
    |||
Db 841 GCGTTCAGCTGCTGCGCGCGCTTCTGTCAGAGCGCGAGGGCGCGCGCGCACGCTGCAAGCG 900
QY 901 CTGTGGAGGCACTCGAGGAGAGAGAGCTCAGCAGCGCTGGCAGAGGACTTGTGCGGCGCTG 960
    |||
Db 901 CTGTGGAGGCACTCGAGGAGAGAGAGCTCAGCAGCGCTGGCAGAGGACTTGTGCGGCGCTG 960
QY 961 ACCGATCCCAATGCGCGCTGCGCTAGACCAAGGGGTGCAAGCCAGCTTTTGAAGAACTTGG 1020
    |||
Db 961 ACCGATCCCAATGCGCGCTGCGCTAGACCAAGGGGTGCAAGCCAGCTTTTGAAGAACTTGG 1020
QY 1021 ATGGCTTACGCTTCTTCTGCGCGCTTATGCTGAACCCCTGTCATTCACGAGGAGCCCTTA 1080
    |||
Db 1021 ATGGCTTACGCTTCTTCTGCGCGCTTATGCTGAACCCCTGTCATTCACGAGGAGCCCTTA 1080
QY 1081 AACTCAGCTTGGCTATCTGCTGAGACCTGCTGGGGGAGAGTTGATGCTTCCCGCAGAG 1140
    |||
Db 1081 AACTCAGCTTGGCTATCTGCTGAGACCTGCTGGGGGAGAGTTGATGCTTCCCGCAGAG 1140
QY 1141 CAGAGACCTGGGGGTGATCATTTGGGGAATTGCTCAGGTACTTGTATAGAGTGGG 1200
    |||
Db 1141 CAGAGACCTGGGGGTGATCATTTGGGGAATTGCTCAGGTACTTGTATAGAGTGGG 1200
QY 1201 GTGGGGGGGAGCTTGTGAGATCAGCTCAGCTTCTCTCCATCCAGAGCGGGGCTTA 1260
    |||
Db 1201 GTGGGGGGGAGCTTGTGAGATCAGCTCAGCTTCTCTCCATCCAGAGCGGGGCTTA 1260
QY 1261 CAGCAGAGCCCTTACAGTTTCACTCAGAGAGCACTTGTGTTGGTCTCTGAGACTTTCAT 1320
    |||
Db 1261 CAGCAGAGCCCTTACAGTTTCACTCAGAGAGCACTTGTGTTGGTCTCTGAGACTTTCAT 1320
QY 1321 CTTGGGTGCTGAGATATGCACTGAAGTAAACAGAAATCAATCTTGGCTGCGCCAGC 1380
    |||

```

```

Db 1321 CTTGGGTGCTGAGATATGCACTGAAGTAAACAGAAATCAATCTTGGCTGCGCCAGC 1380
QY 1381 TCACACTAGCGCTGGAGCCCGGAATGTTAAGCAATGATTAATTAATACAGCG 1435
    |||
Db 1381 TCACACTAGCGCTGGAGCCCGGAATGTTAAGCAATGATTAATTAATACAGCG 1435

```

RESULT 4

LOCUS BC004491 1475 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, clone MGC:11078 IMAGE:3689007, mRNA, complete cds.

ACCESSION BC004491

VERSION BC004491.1 GI:13325366

KEYWORDS MGC.

SOURCE

Organism

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 1475)

AUTHORS

Strausberg R.

TITLE

Direct Submission

JOURNAL

Submitted (12-MAR-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

<http://www.systemsbio.org>

contact: amadan@systemsbio.org

Anup Madan, Rachel Dickhoff, Jessica Fahy, Stephanie Ford, Julia

Greene, Mark Ketteman and Anuradha Madan

FEATURES

source

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 14 Row: m Column: 21

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis, Similarity but not identity to protein.

Location/Qualifiers

1..1475

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:11078 IMAGE:3689007"

/tissue_type="Pancreas, adenocarcinoma"

/clone_id="NIH MGC_39"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

67..1005

/codon_start=1

/product="Unknown (protein for MGC:11078)"

/protein_id="AAH04491.1"

/db_xref="GI:13325367"

/translation="MAAGQNGHEBVGSAVLFESSLDKVVLDNAVHPQKAVYRA

LQAALASGSGPDVLOMLKTHRSDPQILVDFGROPGRFLRAYEGALRALQSS

LAALIAQHSVPLQLELRAGAERLDALDERCISCTIAOQPRLRDELALEALRS

NKCGSAGRGDGVASAPLOPVPVSLSEVAPPPADOTFLPQGPVNVRLSLKD

QOTFARVGLKMRVGSILRGCAALRDPALDSIAYVEBGLYEQAFOLLRRVQAE

GRATLRLVLALEENELTSLAEIDLGLTTPNGSLA"

BASE COUNT 271 a 442 c 504 g 258 t

ORIGIN

Query Match 99.9%; Score 1433.4; DB 9; Length 1475;

Best Local Similarity 99.9%; Pred. No. 4.1e-220;

Matches 1434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTGGGGGTGGGAACCCAGGCCCGCGGCGCGCAGAGGTGAGATGCGAGCTGGG 60

|||||

|||||

Db 19 CTGGCGGGGCTGGGAACCCAGGCCCCCGGCGCCAGAGAGTGATGACACTGGG 78
 QY 61 CAAATATGGGACGAAAGATGGGTGGGAGGAGCATACCTGTTTGTAGAGTCTCGCTGGAC 120
 Db 79 CAAATATGGGACGAAAGATGGGTGGGAGGAGCATACCTGTTTGTAGAGTCTCGCTGGAC 138
 QY 121 AAGATGATCTGTGTGAGATGCTTACGCGACCCCGAGAGAGGTGGAGTGAACAGGGCT 180
 Db 139 AAGATGATCTGTGTGAGATGCTTACGCGACCCCGAGAGAGGTGGAGTGAACAGGGCT 198
 QY 181 CTGACAGGCTGCTTGGAGAGAGCGGGGAGAGCCCGGAGCTGTGCGAGAGTGAAGATC 240
 Db 199 CTGACAGGCTGCTTGGAGAGAGCGGGGAGAGCCCGGAGCTGTGCGAGAGTGAAGATC 258
 QY 241 CACCGGAGCGAGCCGAGCTGATGCTGACCTGGATTTGCGGGGCGGAGCCCTGTGGC 300
 Db 259 CACCGGAGCGAGCCGAGCTGATGCTGACCTGGATTTGCGGGGCGGAGCCCTGTGGC 318
 QY 301 CGCTTCTCCGCGCTTACCGCGAGGGGGCGCTGGCGCGCTGACAGAGAGCTGGAGC 360
 Db 319 CGCTTCTCCGCGCTTACCGCGAGGGGGCGCTGGCGCGCTGACAGAGAGCTGGAGC 378
 QY 361 GCGCGCTGCGCCGAGCATCTGGTGCCTGCACTGAGAGCTGGCGCGCGCGCGCGAGCGG 420
 Db 379 GCGCGCTGCGCCGAGCATCTGGTGCCTGCACTGAGAGCTGGCGCGCGCGCGCGAGCGG 438
 QY 421 CTGGAAGCTTGTGTGGGCGGAGGAGAGCGCTGTTTGAAGTTCATCTTACGCCAGAGAGCC 480
 Db 439 CTGGAAGCTTGTGTGGGCGGAGGAGAGCGCTGTTTGAAGTTCATCTTACGCCAGAGAGCC 498
 QY 481 GACCGGCTCCGCGGAGTGAAGAACTGGCTGAGCTGAGAGAGTGCCTGCGAAATCTTGAATGC 540
 Db 499 GACCGGCTCCGCGGAGTGAAGAACTGGCTGAGCTGAGAGAGTGCCTGCGAAATCTTGAATGC 558
 QY 541 GGTCTGGGGGGGCGGGGGTGGCGAGCGGGGAGGTGCTTCCGCCCTTTCGAGCCCGCGGTG 600
 Db 559 GGTCTGGGGGGGCGGGGGTGGCGAGCGGGGAGGTGCTTCCGCCCTTTCGAGCCCGCGGTG 618
 QY 601 CCTCTCTGTGTGAGAGTGAAGCGCGCGCGCGCGCGAGCTGGCGAGACTTTTCTGTTTC 660
 Db 619 CCTCTCTGTGTGAGAGTGAAGCGCGCGCGCGCGCGAGCTGGCGAGACTTTTCTGTTTC 678
 QY 661 CAGAGGTGAGCTGTAGTGAATCGGCGCGCTGAGCTGAAAGAACCAAGAGCTTTCGCGCGC 720
 Db 679 CAGAGGTGAGCTGTAGTGAATCGGCGCGCTGAGCTGAAAGAACCAAGAGCTTTCGCGCGC 738
 QY 721 TCTGTGGGTCTCAATATGGCGCAAGTGGGGGCTCACTGCAAGCGAGGCTGCCGGGCGCTG 780
 Db 739 TCTGTGGGTCTCAATATGGCGCAAGTGGGGGCTCACTGCAAGCGAGGCTGCCGGGCGCTG 798
 QY 781 CGGAGACCGCGCGCTGAGCTGGGCTTACAGATGACAGGGCGAGGGAGCTGTACGAGCAG 840
 Db 799 CGGAGACCGCGCGCTGAGCTGGGCTTACAGATGACAGGGCGAGGGAGCTGTACGAGCAG 858
 QY 841 GCGTTTCCAGCTGTGCGCGCGCTTCTGTCAGAGCCGAGGGCGCGCGCGCTGCAAGCGC 900
 Db 859 GCGTTTCCAGCTGTGCGCGCGCTTCTGTCAGAGCCGAGGGCGCGCGCGCTGCAAGCGC 918
 QY 901 CTGTGTGAGGAGCTCGAGAGAGAGAGAGCTCAACAGCTTGGCGAGAGACTTGTGGGCGCTG 960
 Db 919 CTGTGTGAGGAGCTCGAGAGAGAGAGAGCTCAACAGCTTGGCGAGAGACTTGTGGGCGCTG 978
 QY 961 ACCGATCCCAATGGCGCGCTGAGCTAGACAGGGGTGACAGCGCTTGTGAGAAACCTGG 1020
 Db 979 ACCGATCCCAATGGCGCGCTGAGCTAGACAGGGGTGACAGCGCTTGTGAGAAACCTGG 1038
 QY 1021 ATGGCTTTAAGGTTCTTCTGCGGCTATGCTGAAACCTGTCCATCCAGGGAGCCTTGA 1080
 Db 1039 ATGGCTTTAAGGTTCTTCTGCGGCTATGCTGAAACCTGTCCATCCAGGGAGCCTTGA 1098
 QY 1081 AACTCCACTTGGCTTATCTGCTGAGACTGCTGGGGCAGAGTTGATTCCTTCCCAAGAG 1140
 Db 1099 AACTCCACTTGGCTTATCTGCTGAGACTGCTGGGGCAGAGTTGATTCCTTCCCAAGAG 1158

QY 1141 CCAGACCACCTGGGGGTGATCATATGGGGATTTGCTCTCAGAGTACTTGTATGAGTGTGGG 1200
 Db 1159 CCAGACCACCTGGGGGTGATCATATGGGGATTTGCTCTCAGAGTACTTGTATGAGTGTGGG 1218
 QY 1201 GTGGGGGGGAGCTTGTGGAGATCAGCCTTCTTCCATCCAGAAACCGGGGCTTGA 1260
 Db 1219 GTGGGGGGGAGCTTGTGGAGATCAGCCTTCTTCCATCCAGAAACCGGGGCTTGA 1278
 QY 1261 CAGCCAGCCCTTACATTCATCATGAAGACCTGATCTTGTGGTGTCTGTGAGACTTCAAT 1320
 Db 1279 CAGCCAGCCCTTACATTCATCATGAAGACCTGATCTTGTGGTGTCTGTGAGACTTCAAT 1338
 QY 1321 CCTGGGTGCTGAGATATCTGAGTGAAGTAAACAGAAATCAATCTTGGCTGCCCCAGC 1380
 Db 1339 CTTGGGTGCTGAGATATCTGAGTGAAGTAAACAGAAATCAATCTTGGCTGCCCCAGC 1398
 QY 1381 TCACACTCAGCTGGGAGCCCGAGTGTAAAGCAATGATTAATAAGTATTAACACGG 1435
 Db 1399 TCACACTCAGCTGGGAGCCCGAGTGTAAAGCAATGATTAATAAGTATTAACACGG 1453

RESULT 5
 AK090673
 LOCUS
 DEFINITION
 AK090673 1833 bp mRNA linear PRI 15-JUL-2002
 Homo sapiens CDNA FLJ33354 fis, clone BRACE2005124, highly similar
 to TUMOR NECROSIS FACTOR RECEPTOR TYPE 1 ASSOCIATED DEATH DOMAIN
 PROTEIN.
 ACCESSION
 AK090673
 VERSION
 AK090673.1 GI:21748880
 KEYWORDS
 oligo capping; fis (full insert sequence).
 SOURCE
 Homo sapiens cerebellum CDNA to mRNA, clone lib:BRACE2
 clone:BRACE2005124.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hirao, M.,
 Shimizu, F., Wakebe, H., Ono, T., Hishigaki, H., Matanabe, T., Ozaki, K.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
 Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
 Kikuchi, H., Kanda, K., Magatsu, M., Murakawa, K., Kanehori, K.,
 Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
 Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
 TITLE
 NEDO human CDNA sequencing project
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 1833)
 AUTHORS
 Isogai, T. and Yamamoto, J.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kametani, Kisekazu, Chiba 292-0812, Japan
 (E-mail:genom@csbri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human CDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; CDNA full insert sequencing:
 Research Association for Biotechnology (RAB) (CDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5' - & 3' - end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.

FEATURES
 source
 1..1833
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="BRACE2005124"
 /tissue_type="cerebellum"
 /clone_lib="BRACE2"
 /note="Cloning vector: pME18SFL3"

BASE COUNT
 ORIGIN
 Query Match 86.7%; Score 1243.8; DB 9; Length 1833;

Best Local Similarity 98.7%; Pred. No. 8.9e-190;
Matches 1254; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

QY 165 GGCAGTGAACAGGAGCTGCAAGCGCTGCTTGACAGAGAGCGCGGAGACCGCGGACCTGTCT 224
Db 561 GGGACCGCGCGCGCTTGAACCGCGCTTCCGACAGAGAGCGCGGAGAGCGGAGCTGTCT 620
QY 225 GCAGATGCTGAAGATCCACCGCAGCGACCGCGAGCTGATCTGTGACGTGCGATTTCGCG 284
Db 621 GCAATGCTGAAGATCCACCGCAGCGACCGCGAGCTGATCTGTGACGTGCGATTTCGCG 680
QY 285 GCGGAGCTCTGTGCGCGCTTCTCTCGCGCTTACCGCGAGGAGGAGCGCTGCGCGCGCT 344
Db 681 GCGGAGCTCTGTGCGCGCTTCTCTCGCGCTTACCGCGAGGAGGAGCGCTGCGCGCGCT 740
QY 345 GCAGAGGAGCGCTGCGCGCGCTGCGCGACGCTGCGGAGCTGCGGAGCTGCGGAGCTGCG 404
Db 741 GCAGAGGAGCGCTGCGCGCGCTGCGCGACGCTGCGGAGCTGCGGAGCTGCGGAGCTGCG 800
QY 405 CGCGCGCGCGCGCGCTGCGCGCTGCGCGAGCGAGCGCGCTGTTGAGTTGAGT 464
Db 801 CGCGCGCGCGCGCGCTGCGCGCTGCGCGAGCGAGCGCGCTGTTGAGTTGAGT 860
QY 465 CCTAGCCCGACGAGCGCGCTGCGCGAGTGAAGACTGCTGAGCTGAGAGAGTGCCT 524
Db 861 CCTAGCCCGACGAGCGCGCTGCGCGAGTGAAGACTGCTGAGCTGAGAGAGTGCCT 920
QY 525 GCGAATCTGAAGAGCGCGCTGCGCGCGCGCGCGCGAGGAGTGCCTGCGCGCG 584
Db 921 GCGAATCTGAAGAGCGCGCTGCGCGCGCGCGCGCGAGGAGTGCCTGCGCGCG 980
QY 585 CTGCGAGCGCGCGCGCTGCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCACTGTC 644
Db 981 CTGCGAGCGCGCGCGCTGCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCACTGTC 1040
QY 645 CCAGACTTTTCTGTTCCAGAGGTGACCTTGAATCGGCGCGCTGAGCTGAAGAGCA 704
Db 1041 CCAGACTTTTCTGTTCCAGAGGTGACCTTGAATCGGCGCGCTGAGCTGAAGAGCA 1100
QY 705 ACAGAGCTTCTGCGCGCTGCTGAGGTCTCAAAAGGCGCAAGTGGGCGCTGCACTGAGCG 764
Db 1101 ACAGAGCTTCTGCGCGCTGCTGAGGTCTCAAAAGGCGCAAGTGGGCGCTGCACTGAGCG 1160
QY 765 AGGCTGCGCGCGCTGCGCGAGCGCGCGCTGAGCTGCTGAGCTGAGTGAAGAGCGCA 824
Db 1161 AGGCTGCGCGCGCTGCGCGAGCGCGCGCTGAGCTGCTGAGCTGAGTGAAGAGCGCA 1220
QY 825 GGGACTGTACGAGCAGGCGCTTCCAGCTGCTGCGCGCTTCTGTCAGAGCGGAGCGCGCG 884
Db 1221 GGGACTGTACGAGCAGGCGCTTCCAGCTGCTGCGCGCTTCTGTCAGAGCGGAGCGCGCG 1280
QY 885 GCGCAGCGCTGCGCGCGCTGCTGAGGAGCTGAGGAGAGAGAGTCAACAGCGCTGCGAGA 944
Db 1281 GCGCAGCGCTGCGCGCGCTGCTGAGGAGCTGAGGAGAGAGAGTCAACAGCGCTGCGAGA 1340
QY 945 GGAATTGCTGCGCGCTGAGCGAGCTTCCAGCTGCTGCGCGCTTCCAGAGCGGAGTGA 1004
Db 1341 GGAATTGCTGCGCGCTGAGCGAGCTTCCAGCTGCTGCGCGCTTCCAGAGCGGAGTGA 1400
QY 1005 CTTTGGAGAACTGTGATGCGCTTAAAGTTCTTCTGCGCGCTTCTGTCAGACCCCTGTCC 1064
Db 1401 CTTTGGAGAACTGTGATGCGCTTAAAGTTCTTCTGCGCGCTTCTGTCAGACCCCTGTCC 1460
QY 1065 ATCCACGCGGAGCTGAGAACTCCACCTTGGGCTATCTGCTGAGACCTGCTGAGGAGTGA 1124
Db 1461 ATCCACGCGGAGCTGAGAACTCCACCTTGGGCTATCTGCTGAGACCTGCTGAGGAGTGA 1520
QY 1125 TTGCTTCCCGAGAGCGCAGACCACTGCGGAGTGCATCATTTGGAGATTCTGCTCAGATAC 1184
Db 1521 TTGCTTCCCGAGAGCGCAGACCACTGCGGAGTGCATCATTTGGAGATTCTGCTCAGATAC 1580
QY 1185 TTGATAGATGTGGGTGGGGGAGCTTGCTTTGAGATCAGCTCAGCTTCTCCATC 1244

```

```

Db 1581 TTTATAGACTGTGGGTGGGGGAGCTGCTTTGAGATCAGCTCAGCTTCTCCATC 1640
QY 1245 CCAGAGCGGCGCTTACAGCGAGCTTACAGTTTCACTGATGAAGACCTTGTCTT 1304
Db 1641 CCAGAGCGGCGCTTACAGCGAGCTTACAGTTTCACTGATGAAGACCTTGTCTT 1700
QY 1305 GTGTCTGGAATTTATCTGTGTCTGTCAGATATCTGAGTGAAGTAAACAGGAATCAAT 1364
Db 1701 GTGTCTGGAATTTATCTGTGTCTGTCAGATATCTGAGTGAAGTAAACAGGAATCAAT 1760
QY 1365 CTTGCTGCGCGCGCGAGCTCAGACTGAGCGTGGAGACCGGAATGTTAAGCAATGATATAA 1424
Db 1761 CTTGCTGCGCGCGAGCTCAGACTGAGCGTGGAGACCGGAATGTTAAGCAATGATATAA 1820
QY 1425 GTATACACGG 1435
Db 1821 GTATACACGG 1831

RESULT 6
AC074143/c 152953 bp DNA linear PRI 01-MAY-2002
LOCUS AC074143
DEFINITION Homo sapiens chromosome 16 clone RP11-5A19, complete sequence.
ACCESSION AC074143
VERSION AC074143.4 GI:20376965
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 152953)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 152953)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Direct Submission
REFERENCE 3 (bases 1 to 152953)
AUTHORS Submitted (01-MAY-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
Source
Location/Qualifiers
1..152953
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-5A19"

BASE COUNT 38753 a 35341 c 36249 g 42610 t
ORIGIN

```


Query Match 70.1%; Score 1005.8; DB 9; Length 152953;
 Best Local Similarity 84.2%; Pred. No. 5,7e-152;
 Matches 1254; Conservative 0; Mismatches 17; Indels 218; Gaps 2;

QY 165 GGCAGGTACAGGGCTCTGACAGGCTGCTTGGCAGAGACCGCGGAGACCCTGGACGTGCT 224
 Db 124325 GGGAGCCCGCCCGCCCTGACCGCCCGCTCCGACAGAGCGCGCGAGCCCGGACGTGCT 124266

QY 225 GCAGATGCTGAAGATCCACCGCAGACCGCCAGCTGATGCTGAGCTGGAGTCTGCGG 284
 Db 124265 GCAGATGCTGAAGATCCACCGCAGACCGCCAGCTGATGCTGAGCTGGAGTCTGCGG 124206

QY 285 GCGCAGACCCCTGTGCGCGCTTCTCCGCGCTTACCGCGAGGGGCGCTGCGCGCGCT 344
 Db 124205 GCGCAGACCCCTGTGCGCGCTTCTCCGCGCTTACCGCGAGGGGCGCTGCGCGCGCT 124146

QY 345 GCAGAGAGCTTGGCGCGCGCGCTGCGCGACACTCGCTGCGCACTGAGCTGCG 404
 Db 124145 GCAGAGAGCTTGGCGCGCGCGCTGCGCGACACTCGCTGCGCACTGAGCTGCG 124086

QY 405 CGCGCGCGCGCGCGCGCGCTTGTGCGGAGAGAGGCGCTGTTGAGTTCAT 464
 Db 124085 CGCGCGCGCGCGCGCGCTTGTGCGGAGAGAGGCGCTGTTGAGTTCAT 124026

QY 465 CCTAGCCGACG-----AGCCCGACCGGCTCCGCGAGTAAGAAC 475
 Db 124025 CCTAGCCGACGAGTGGCGCGCGCTAGGCTAGGATGGGGTGGCGGAGTCCGCGCT 123966

QY 476 -----AGCCCGACCGGCTCCGCGAGTAAGAAC 502
 Db 123965 TAAATCCCCCGTGACGCGACGTTCCATTTCCCTAGCCGACCGGCTCCGCGAGTAAGAAC 123906

QY 503 TGGCTGAGCTGAGAGATGCGCTGCGAATCTGAGTGGGTGCGGGGCGCGGGGTGCG 562
 Db 123905 TGGCTGAGCTGAGAGATGCGCTGCGAATCTGAGTGGGTGCGGGGCGCGGGGTGCG 123846

QY 563 ACCGGGAGGTCGCTTCCGCGCGCTTGGAGCGCGCGTCCCTCTGTCGAGAGTGAAGC 622
 Db 123845 ACCGGGAGGTCGCTTCCGCGCGCTTGGAGCGCGCGTCCCTCTGTCGAGAGTGAAGC 123786

QY 623 CGCGCGCGCGCGCGCACTGCGCACTTTTCTGTTCCAGGGTCAAGCTGT----- 674
 Db 123785 CGCGCGCGCGCGCGCACTGCGCACTTTTCTGTTCCAGGGTCAAGCTGTAGTGAGG 123726

QY 675 ----- 674
 Db 123725 CGCAGGTGAGAGCGCGGGTCAACAGAGATTGAGTCTGGGAGGGGCAAGGGTCTGTG 123666

QY 675 ----- 674
 Db 123665 CCGGGTGGGAGTGGGTGGACCGGCGCTGGGGGTCCCGCCCTTACCCCGTTTCCCACT 123606

QY 675 -----AGTAATCGGCGCGCTGAGCTTGAAGCAACAGAGCTTCCGCGCTGTG 726
 Db 123605 CCTCTGCAAGTGAATCGGCGCGCTGAGCTTGAAGCAACAGAGCTTCCGCGCTGTG 123546

QY 727 GGTTCGAATGCGCAGAGTGGGGCGCTCACTGAGAGAGGCTGCGGGCGCTTCCGGGAC 786
 Db 123545 GGTTCGAATGCGCAGAGTGGGGCGCTCACTGAGAGAGGCTGCGGGCGCTTCCGGGAC 123486

QY 787 CCGCGCTGAGCTGCTGCTGCTTGAAGTGAAGAGCGGAGGAGCTGTGAGCAGAGCTTC 846
 Db 123485 CCGCGCTGAGCTGCTGCTGCTTGAAGTGAAGAGCGGAGGAGCTGTGAGCAGAGCTTC 123426

QY 847 CAGCTGCTGCGCGCTTGTGTCAGGCGCGGCGCGCGCTGCGAGCTGCGAGCTGCTG 906
 Db 123425 CAGCTGCTGCGCGCTTGTGTCAGGCGCGGCGCGCGCTGCGAGCTGCGAGCTGCTG 123366

QY 907 GAGGCACTCGAGAGAGAGAGCTCAACAGCTGCGAGAGACTTGGCGCTTGAACCGAT 966
 Db 123365 GAGGCACTCGAGAGAGAGAGCTCAACAGCTGCGAGAGACTTGGCGCTTGAACCGAT 123306

QY 967 CCNATGCGGCGCTGCGCTTAGACAGGGGTGACGACGCTTTTGGAGAACCTGATGCGC 1026
 Db 123305 CCNATGCGGCGCTGCGCTTAGACAGGGGTGACGACGCTTTTGGAGAACCTGATGCGC 123246

QY 1027 TTAGGGTCTTCTTGGCGCTTATGCTGAGACCCCTGTCATCCAGGGACCTTGAACCTCC 1086
 Db 123245 TTAGGGTCTTCTTGGCGCTTATGCTGAGACCCCTGTCATCCAGGGACCTTGAACCTCC 123186

QY 1087 ACTTGGCTTCTGCTGAGACCTGCTGGGCGAGAGTGTGATGCTTCCCGAGAGCCAGAC 1146
 Db 123185 ACTTGGCTTCTGCTGAGACCTGCTGGGCGAGAGTGTGATGCTTCCCGAGAGCCAGAC 123126

QY 1147 CACTGGGGGTGATCATTTGGGGAATTCCTCCTCAGTAATTGTAAGTGTGGGATGGG 1206
 Db 123125 CACTGGGGGTGATCATTTGGGGAATTCCTCCTCAGTAATTGTAAGTGTGGGATGGG 123066

QY 1207 GGGACTTGTCTTGGAGATCAGCTTCTTCCATCCCAAGGGGGCTTACAGCCA 1266
 Db 123065 GGGACTTGTCTTGGAGATCAGCTTCTTCCATCCCAAGGGGGCTTACAGCCA 123006

QY 1267 GCCCTTACAGTTTCACTCATGAGACCTTATCTTGGTGTCTGAGACTTCACTCCCTGGG 1326
 Db 123005 GCCCTTACAGTTTCACTCATGAGACCTTATCTTGGTGTCTGAGACTTCACTCCCTGGG 122946

QY 1327 TGTGCAGATACCTGCAAGTAAAGTAAACAGAAATCAATCTTGGCTGCCCGCAGCTCAC 1386
 Db 122945 TGTGCAGATACCTGCAAGTAAAGTAAACAGAAATCAATCTTGGCTGCCCGCAGCTCAC 122886

QY 1387 TCAAGCTGGGACCCCGAATGTTAAGCAATGATTAATTAAGTAAACAGG 1435
 Db 122885 TCAAGCTGGGACCCCGAATGTTAAGCAATGATTAATTAAGTAAACAGG 122837

RESULT 7
 AC040160 209574 bp DNA linear PRI 31-MAR-2001
 LOCUS Homo sapiens chromosome 16 clone CTC-277H1, complete sequence.
 DEFINITION AC040160
 ACCESSION AC040160.4 GI:13491214
 VERSION
 KEYWORDS HTG.
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 209574)
 DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE
 2 (bases 1 to 209574)
 DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (31-MAR-2001) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 On Mar 31, 2001 this sequence version replaced gi:9280711.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 www.sngc.stanford.edu
 Quality: Phrap Quality >=40 99.7% of Sequence;
 Estimated Total Number of Errors is 0.5.
 SRS Content:
 SHGC-11402 G14546.

FEATURES
 source
 1..209574
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="CTC-277H1"

BASE COUNT 50328 a 55727 c 55548 g 47971 t
ORIGIN
Query Match 70.1%; Score 1005.8; DB 9; Length 209574;
Best Local Similarity 84.2%; Pred. No. 5.4e-152;
Matches 1254; Conservative 0; Mismatches 17; Indels 218; Gaps 2;
QY 165 GGCAGTGTACAGGGCTCTGTCAGAGCTGCTTGTGACAGAGCGGGAGCCCGAGCTGCT 224
Db 183123 GGGACCGCGCCGCGCTGACCGCCGCTCCGAGAGAGCGGGAGCCCGAGCTGCT 183182
QY 225 GCAAGTGTGAAAGATTCACCGGAGAGGAGCCCGAGAGTGTGACAGTCTGCTGCGG 284
Db 183183 GCAAGTGTGAAAGATTCACCGGAGAGGAGCCCGAGAGTGTGACAGTCTGCTGCGG 183242
QY 285 GGGGAGAGCTGTGAGCGCTTCTCTCGCGCTACCGGAGGGGGCGCTGCGCGCGCT 344
Db 183243 GCGGAGAGCTGTGAGCGCTTCTCTCGCGCTACCGGAGGGGGCGCTGCGCGCGCT 183302
QY 345 GCAGAGAGAGCTGTGAGCGCGCGCTGCGCGAGCACTGTGAGTGTGCAACTGAGAGCTGCG 404
Db 183303 GCAGAGAGAGCTGTGAGCGCGCGCTGCGCGAGCACTGTGAGTGTGCAACTGAGAGCTGCG 183362
QY 405 CGCCGCGCGCGCGAGCGCTGTGAGAGCTTGTGCGGAGAGAGCGCTGTTTGAAGTGTGAT 464
Db 183363 CGCCGCGCGCGAGCGCTGTGAGAGCTTGTGCGGAGAGAGCGCTGTTTGAAGTGTGAT 183422
QY 465 CCTAGCGCGAGC----- 475
Db 183422 CTTAGCGCGAGAGCTGTGAGCGCGCGAGCTAGAGATGGGGATAGGGCGGGATCCGCGGT 183482
QY 476 -----AGCCGAGCGGCTCCGAGATGAAGAAC 502
Db 183483 TAAGTCCCGCGTGTGAGCGAGCACTTCATTCCTCAGCCGAGCTCCGAGATGAAGAAC 183542
QY 503 TGGCTGAGCTGTGAGAGATGCGCTGTGCGAAATCTGAATGTGCGGCTTCGGGGGCGCGGGGTGCGG 562
Db 183543 TGGCTGAGCTGTGAGAGATGCGCTGTGCGAAATCTGAATGTGCGGCTTCGGGGGCGCGGGGTGCGG 183602
QY 563 AGCGGGAGAGTCTGCTTGTGCGCGCGCGCTTGTGCGAGCGCGCGCTTGTGCGAGAGTGAAC 622
Db 183603 AGCGGGAGAGTCTGCTTGTGCGCGCGCGCTTGTGCGAGCGCGCGCTTGTGCGAGAGTGAAC 183662
QY 623 CGCGCGCGCGCGCGCGAGCTGTGCGCGAGACTTCTGTTCAGAGGTGAGCGCTGT----- 674
Db 183663 CGCGCGCGCGCGCGCGAGCTGTGCGCGAGACTTCTGTTCAGAGGTGAGCGCTGTGAGAGT 183722
QY 675 ----- 674
Db 183723 CGCAGGTGTGAGAGCGCGGGGTCAACAGAGAGTGTGAGTGTGGGGAGGGGAGGAGGTGCTTG 183782
QY 675 ----- 674
Db 183783 CCGGGTGGGAGATGGGTGGAGCCGCGCGCTGGGGGTCCCGCGCTCACCCTTTTCCCACT 183842
QY 675 -----AGTGAATGTGCGCGCTGTGAGCTGTGAAGAGCAACAAGCTTTCGCGCTGTGTG 726
Db 183843 CCCTGTGAGTGTGATGCGCGGTGAGCTGTGAAGAGCAACAAGCTTTCGCGCTGTGTG 183902
QY 727 GGTCTCAAAATGTGCGCAAGGTGGGGCGCTCACTGTGAGAGAGCTGTGCGGGCGCTGTGCGGAGC 786
Db 183903 GGTCTCAAAATGTGCGCAAGGTGGGGCGCTCACTGTGAGAGAGCTGTGCGGGCGCTGTGCGGAGC 183962
QY 787 CGGGGCGCTGTGAGCTGTGAGCTGTGAGTGTGAGAGAGAGAGAGTGTGAGAGAGAGAGCTTTC 846
Db 183963 CGGGGCGCTGTGAGCTGTGAGCTGTGAGTGTGAGAGAGAGAGAGTGTGAGAGAGAGAGCTTTC 184022
QY 847 CAGCTGTCTGTGAGAGCTGTGAGAGCGCGAGCGCGCGCGCACTGTGAGAGCGCTGTGAGTGTG 906
Db 184023 CAGCTGTCTGTGAGAGCTGTGAGAGCGCGAGCGCGCGCGCACTGTGAGAGCGCTGTGAGTGTG 184082
QY 907 GAGGCACTGTGAGAGAGAGAGAGCTCACAGAGCTGTGAGAGAGAGAGCTTGTGAGAGAGTGTG 966

Db 184083 GAGGCACTGTGAGAGAGAGAGAGCTCACAGAGCTGTGAGAGAGAGCTTGTGAGAGAGTGTG 184142
QY 967 CCAATGTGCGCGCTGTGAGCTGTGAG 1026
Db 184143 CCAATGTGCGCGCTGTGAGCTGTGAG 184202
QY 1027 TTAGGGTCTTCTGTGCGGCTGTGAG 1086
Db 184203 TTAGGGTCTTCTGTGCGGCTGTGAG 184262
QY 1087 ACTTGGCTATCTGTGTGAG 1146
Db 184263 ACTTGGCTATCTGTGTGAG 184322
QY 1147 CACTGGGGGTGCATCATTTGGAGATTTGCTGCTCAGATCATTTGATGATGATGATGATGATGATGAT 1206
Db 184323 CACTGGGGGTGCATCATTTGGAGATTTGCTGCTCAGATCATTTGATGATGATGATGATGATGATGATGAT 184382
QY 1207 GGGAGCTTCTTGTGAGATCAGCTTCTTCCATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1266
Db 184383 GGGAGCTTCTTGTGAGATCAGCTTCTTCCATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 184442
QY 1267 GCGCTTACAGTTTCACTGATGAG 1326
Db 184443 GCGCTTACAGTTTCACTGATGAG 184502
QY 1327 TGCTGTGAGATCTGTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1386
Db 184503 TGCTGTGAGATCTGTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 184562
QY 1387 TCAGCGTGTGAG 1435
Db 184563 TCAGCGTGTGAG 184611

RESULT 8
AC008594/c
LOCUS 228516 bp DNA linear HTG 21-JUN-2000
DEFINITION Homo sapiens chromosome 16 clone CTA-277H1, WORKING DRAFT SEQUENCE,
34 unordered pieces.
AC008594
AC008594.6 GI:8575891
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 228516)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 228516)
REFERENCE DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT
On Jun 21, 2000 this sequence version replaced gi:7689774.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1258007
Center clone name: C19788KA_277H1

Summary Statistics
Consensus quality: 197894 bases at least Q40
Consensus quality: 211598 bases at least Q30
Consensus quality: 214743 bases at least Q20
Estimated insert size: 212740; agarose-fp estimation
Estimated insert size: 225216; sum-of-ctnigs estimation
Quality coverage: 6.02 in Q20 bases; agarose-fp estimation

Quality coverage: 5.68 in Q20 bases; sum-of-contigs estimation.

* NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 1001: contig of 1001 bp in length
1002 1101: gap of unknown length
1102 2101: contig of 1000 bp in length
2102 2201: gap of unknown length
2202 3354: contig of 1153 bp in length
3355 3455: gap of unknown length
3456 4556: contig of 1102 bp in length
4557 4656: gap of unknown length
4657 5682: contig of 1026 bp in length
5683 5782: gap of unknown length
5783 7157: contig of 1375 bp in length
7158 7257: gap of unknown length
7258 8361: contig of 1104 bp in length
8362 8461: gap of unknown length
8462 9465: contig of 1004 bp in length
9466 9565: gap of unknown length
9566 10598: contig of 1033 bp in length
10599 11799: gap of unknown length
11800 11899: contig of 1101 bp in length
11900 13061: gap of unknown length
13062 13161: contig of 1162 bp in length
13162 14164: contig of 1003 bp in length
14165 14264: gap of unknown length
14265 16006: contig of 1742 bp in length
16007 16106: gap of unknown length
16107 18072: contig of 1966 bp in length
18073 18172: gap of unknown length
18173 20504: contig of 2332 bp in length
20505 20604: gap of unknown length
20605 24399: contig of 3795 bp in length
24400 24499: gap of unknown length
24500 28638: contig of 4139 bp in length
28639 32855: contig of 4117 bp in length
32856 32955: gap of unknown length
32956 35940: contig of 2985 bp in length
35941 36040: gap of unknown length
36041 39489: contig of 3449 bp in length
39490 39589: gap of unknown length
39590 41280: contig of 1691 bp in length
41281 41380: gap of unknown length
41381 49404: contig of 8024 bp in length
49405 49504: gap of unknown length
49505 53208: contig of 3704 bp in length
53209 53308: gap of unknown length
53309 58096: contig of 4788 bp in length
58097 58196: gap of unknown length
58197 65406: contig of 7210 bp in length
65407 65506: gap of unknown length
65507 73598: contig of 8092 bp in length
73599 73698: gap of unknown length
73699 79746: contig of 6048 bp in length
79747 79846: gap of unknown length
79847 89098: contig of 9252 bp in length
89099 89198: gap of unknown length
89199 98103: contig of 8805 bp in length
98104 98105: gap of unknown length
98106 108051: contig of 9948 bp in length
108052 108151: gap of unknown length
108152 126494: contig of 18343 bp in length
126495 126594: gap of unknown length
126595 149165: contig of 22571 bp in length
149166 149265: gap of unknown length
```

* 149266 177969: contig of 28704 bp in length
* 177970 178069: gap of unknown length
* 178070 228516: contig of 50447 bp in length.

FEATURES

source

1..228516
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone_id="CTA-277H1"

BASE COUNT 53125 a 60116 c 58426 g 53336 t 3513 others
ORIGIN

Query Match 69.0%; Score 990.8; DB 2; Length 228516;
Best local similarity 84.0%; Pred. No. 1.3e-149;
Matches 1253; Conservative 0; Mismatches 17; Indels 222; Gaps 3;

```
QY 165 GGCAGTGTACAGGCTCTGAGGCTCTTGGCAGAGAGCGGAGCCCGACGTGCT 224
Db 168541 GGGACCGCCCGCGCTGACCGCCGCTCCGACAGAGCGGAGCCCGACGTGCT 168482
QY 225 GCAGATGCTGAAGATCCACCGCAGCAGCCGACGCTGATGTCAGCTGCGATTTGCGG 284
Db 168481 GCAGATGCTGAAGATCCACCGCAGCAGCCGACGCTGATGTCAGCTGCGATTTGCGG 168422
QY 285 GCGGACGCCCTGTGGCGCTTCTCCGCGCTTACCGCGAGGCGGCGCTGCGCGCGCT 344
Db 168421 GCGGACGCCCTGTGGCGCTTCTCCGCGCTTACCGCGAGGCGGCGCTGCGCGCGCT 168362
QY 345 GCAGAGAGAGCTGCGCGCGCGCTGCGCGCAGCACTGCGTCCGCTCAACTGAGAGCTGCG 404
Db 168361 GCAGAGAGAGCTGCGCGCGCGCTGCGCGCAGCACTGCGTCCGCTCAACTGAGAGCTGCG 168302
QY 405 CGCGGCGCGCGAGCGGCTGAGCGCTTCTGCGGCGAGAGAGCGCTGTTGAGTTGCAAT 464
Db 168301 CGCGGCGCGCGAGCGGCTGAGCGCTTCTGCGGCGAGAGAGCGCTGTTGAGTTGCAAT 168242
QY 465 CCGAGCGCCAGC----- 475
Db 168241 CTTAGCCACAGAGCTGCGCGCGGCTAGAGGTAGAGTGGGGTGGGAGGATCCCGGCT 168182
QY 476 -----AGCCGACCGCGCTCCGGAGTGAAGAAC 502
Db 168181 TAAGTCCCCGTGAGCGGCAAGTTCCATTCCCTAGGCCGACCGGCTCCGGAGTGAAGAAC 168122
QY 503 TGGCTAGCTGAGAGATGCGCTGCGCAATCTGAAGTGCAGCTTCCGAGGTGCGG 562
Db 168121 TGGCTAGCTGAGAGATGCGCTGCGCAATCTGAAGTGCAGCTTCCGAGGTGCGG 168062
QY 563 ACGGGAGAGTGGCTTGGGCGCCCTTGCAGACCCCGGTCGCTCTGTGTGAGGTGAAC 622
Db 168061 ACGGGAGAGTGGCTTGGGCGCCCTTGCAGACCCCGGTCGCTCTGTGTGAGGTGAAC 168002
QY 623 CGCGCGCGCGCGCGCACCTGCGCAGACTTTTCTGTCCAGGGTCAAGCTGT----- 674
Db 168001 CGCGCGCGCGCGCGCACCTGCGCAGACTTTTCTGTCCAGGGTCAAGCTGTAGTGAAG 167942
QY 675 ----- 674
Db 167941 CGCGAGTGAAGAGCGGGGTATCCCGAGGTGAGATCTGGGAGAGGCGGCGCTCGT 167882
QY 675 ----- 674
Db 167881 TGGCGGTGATGATGAGTGGAGACCGGCGCTGGGGGTCCCGCGCTCACCCGTTTCCCC 167822
QY 675 -----AGTGAATCGGCGCTGAGCTGAAGAGCAACAGACGTTTCGCGCGCTCT 723
Db 167821 ACTCCCTGAGTGAATGAGGCGGCTGAGCCTGAAGAGACC-ACAAGCGTTCCGCGCTCT 167763
QY 724 GTGGGTCTCAAAATGGCGCAAGGTGGGCGCTCACTGACAGAGGCTGCCGGCGCTGCGG 783
Db 167762 GTGGGTCTCAAAATGGCGCAAGGTGGGCGCTCACTGACAGAGGCTGCCGGCGCTGCGG 167703
```


TITLE
JOURNAL
COMMENT

Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehocky,J., Lieu,C., Locke,R., Macdonald,P., Margulis,N.,
McBwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Strange-Thomson,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teafaye,S., Tixrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Submitted (23-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced g1:8225265.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project Information
Center project name: L2764

Center clone name: 3_K_18

NOTE: This record contains 229 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
698 797: gap of 100 bp in length
798 1497: contig of 700 bp in length
1498 1597: gap of 100 bp in length
1598 2265: contig of 668 bp in length
2266 2365: gap of 100 bp in length
2366 3044: contig of 679 bp in length
3045 3144: gap of 100 bp in length
3145 3839: contig of 695 bp in length
3840 3939: gap of 100 bp in length
3940 4642: contig of 703 bp in length
4643 4742: gap of 100 bp in length
4743 5446: contig of 704 bp in length
5447 5546: gap of 100 bp in length
5547 6268: contig of 722 bp in length
6269 6368: gap of 100 bp in length
6369 7066: contig of 698 bp in length
7067 7166: gap of 100 bp in length
7167 7856: contig of 690 bp in length
7857 7956: gap of 100 bp in length
7957 8640: contig of 684 bp in length
8641 8740: gap of 100 bp in length
8741 9415: contig of 675 bp in length
9416 9515: gap of 100 bp in length
9516 10181: contig of 666 bp in length
10182 10281: gap of 100 bp in length
10282 10968: contig of 688 bp in length
10970 11069: gap of 100 bp in length
11070 11760: contig of 691 bp in length
11761 11860: gap of 100 bp in length
11861 12565: contig of 705 bp in length
12566 12665: gap of 100 bp in length
12666 13372: contig of 707 bp in length
13373 13472: gap of 100 bp in length
13473 14162: contig of 690 bp in length
14163 14262: gap of 100 bp in length
14263 14949: contig of 687 bp in length

14950 15049: gap of 100 bp in length
15050 15750: contig of 701 bp in length
15751 15850: gap of 100 bp in length
15851 16524: contig of 674 bp in length
16525 16624: gap of 100 bp in length
16625 17329: contig of 705 bp in length
17330 17429: gap of 100 bp in length
17430 18125: contig of 696 bp in length
18126 18225: gap of 100 bp in length
18226 18911: contig of 686 bp in length
18912 19011: gap of 100 bp in length
19012 19698: contig of 687 bp in length
19699 19798: gap of 100 bp in length
19799 20498: contig of 700 bp in length
20499 20598: gap of 100 bp in length
20599 21313: contig of 715 bp in length
21314 21413: gap of 100 bp in length
21414 22130: contig of 717 bp in length
22131 22230: gap of 100 bp in length
22231 22917: contig of 687 bp in length
22918 23017: gap of 100 bp in length
23018 23713: contig of 696 bp in length
23714 23813: gap of 100 bp in length
23814 24511: contig of 698 bp in length
24512 24611: gap of 100 bp in length
24612 25296: contig of 685 bp in length
25297 25396: gap of 100 bp in length
25397 26074: contig of 678 bp in length
26075 26174: gap of 100 bp in length
26175 26796: contig of 622 bp in length
26797 26896: gap of 100 bp in length
26897 27596: contig of 700 bp in length
27597 27696: gap of 100 bp in length
27697 28400: contig of 704 bp in length
28401 28500: gap of 100 bp in length
28501 29220: contig of 720 bp in length
29221 29320: gap of 100 bp in length
29321 30032: contig of 712 bp in length
30033 30132: gap of 100 bp in length
30133 30825: contig of 693 bp in length
30826 30925: gap of 100 bp in length
30926 31627: contig of 702 bp in length
31628 31727: gap of 100 bp in length
31728 32408: contig of 681 bp in length
32409 32508: gap of 100 bp in length
32509 33192: contig of 684 bp in length
33193 33292: gap of 100 bp in length
33293 33968: contig of 676 bp in length
33969 34068: gap of 100 bp in length
34069 34744: contig of 676 bp in length
34745 34844: gap of 100 bp in length
34845 35526: contig of 682 bp in length
35527 35626: gap of 100 bp in length
35627 36322: contig of 696 bp in length
36323 36422: gap of 100 bp in length
36423 37118: contig of 696 bp in length
37119 37218: gap of 100 bp in length
37219 37946: contig of 728 bp in length
37947 38046: gap of 100 bp in length
38047 38736: contig of 690 bp in length
38737 38836: gap of 100 bp in length
38837 39531: contig of 695 bp in length
39532 39631: gap of 100 bp in length
39632 40318: contig of 687 bp in length
40319 40418: gap of 100 bp in length
40419 41109: contig of 691 bp in length
41110 41209: gap of 100 bp in length
41210 41900: contig of 691 bp in length
41901 42000: gap of 100 bp in length
42001 42619: contig of 619 bp in length
42620 42719: gap of 100 bp in length
42720 43400: contig of 681 bp in length
43401 43500: gap of 100 bp in length

* 43501 44116: contig of 616 bp in length
* 44217 44216: gap of 100 bp
* 44217 44916: contig of 700 bp in length
* 44917 45016: gap of 100 bp
* 45017 45713: contig of 697 bp in length
* 45714 45813: gap of 100 bp
* 45814 46497: contig of 684 bp in length
* 46498 46597: gap of 100 bp
* 46598 47293: contig of 696 bp in length
* 47294 47393: gap of 100 bp
* 47394 48098: contig of 705 bp in length
* 48099 48198: gap of 100 bp
* 48199 48892: contig of 694 bp in length
* 48893 48992: gap of 100 bp
* 48993 49666: contig of 674 bp in length
* 49667 49766: gap of 100 bp
* 49767 50453: contig of 687 bp in length
* 50454 50553: gap of 100 bp
* 50554 51229: contig of 676 bp in length
* 51230 51329: gap of 100 bp
* 51330 52017: contig of 688 bp in length
* 52018 52117: gap of 100 bp
* 52118 52803: contig of 686 bp in length
* 52804 52903: gap of 100 bp
* 52904 53610: contig of 707 bp in length
* 53611 53710: gap of 100 bp
* 53711 54406: contig of 696 bp in length
* 54407 54506: gap of 100 bp
* 54507 55214: contig of 708 bp in length
* 55215 55314: gap of 100 bp
* 55315 56029: contig of 715 bp in length
* 56030 56129: gap of 100 bp
* 56130 56834: contig of 705 bp in length
* 56835 56934: gap of 100 bp
* 56935 57626: contig of 692 bp in length

Query Match 30.3%; Score 434.4; DB 2; Length 208936;
Best Local Similarity 94.0%; Pred. No. 1.7e-60;
Matches 473; Conservative 0; Mismatches 27; Indels 3; Gaps 2;

QY 933 CAGCTGGCAGAGGAGATTGCTGGGCTGACCCGATCCCATGGGCGCTGGCCCTAGACCCAG 992
Db 110442 CAGCTTGGCAAGGCGCTTCTGCCCCC-GCCCCATCCCAAGGGGGCTGGCCCTAGACCCAG 110385
QY 993 GAGTGCAGCAGCTTTTGGAGAACTGAGTGGCTTAGGCTTCTCTGGGCTATTGCT 1052
Db 110384 GGGTGCAGCAGC-TTTGGAGAACTGGGATGGCTTAGGCTTCTTAAGGGGCAAGGCT 110326
QY 1053 GAACCCCTGTCCATCCACGGGAGCCCTGAACCTGCACTTGGCTTATGCTGCACTGCTG 1112
Db 110325 GAACCCCTGTCCAAACCCGGGGAGCCCTGAACCTGCACTTGGCTTATGCTGCACTGCTG 110266
QY 1113 GGGCAGAGTTGATTGCTTCCCAAGAGCCAGACCACTGGGGGTGCATCTTGGGGATTTC 1172
Db 110265 GGGCAGAGTTGATTGCTTCCCAAGAGCCAGACCACTGGGGGTGCATCTTGGGGATTTC 110206
QY 1173 TGCTCAGGTAATTGATTGAGTGTGGGGGTGGGGGGGAACTTGCTTTGGAATGAGCTCA 1232
Db 110205 TGCTCAGGTAATTGATTGAGTGTGGGGGTGGGGGGGAACTTGCTTTGGAATGAGCTCA 110146
QY 1233 CCTTCTCCATCCCAAGAGCGGGCTTAAAGCCAGCCCTTACAGTTTCACTCAATGAAGA 1292
Db 110145 CCTTCTCCATCCCAAGAGCGGGCTTAAAGCCAGCCCTTACAGTTTCACTCAATGAAGA 110086
QY 1293 CCTTGATCTTTGGTGTCTGAGCTTATCTCTGGGTCTGCAATCTGCAATGAATGATAA 1352
Db 110085 CCTTGATCTTTGGTGTCTGAGCTTATCTCTGGGTCTGCAATCTGCAATGAATGATAA 110026
QY 1353 ACAGGAATCAATCTTGCGTGGCCCCGAGCTCACACTCAGGGTGGGAGCCCGGAATGTTAAAC 1412
Db 110025 ACAGGAATCAATCTTGCGTGGCCCCGAGCTCACACTCAGGGTGGGAGCCCGGAATGTTAAAC 109966
QY 1413 AATGATATAAAGTATACACGG 1435

Db 109965 AATGATATAAAGTATACACGG 109943

|||||
RESULT 11
LOCUS 127062 1384 bp DNA linear PAT 06-FEB-1997
DEFINITION Sequence 3 from patent US 5563039.
ACCESSION 127062
VERSION 127062.1 GI:1817838
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1384)
AUTHORS Goeddel, D. V. and Hsu, H.
TITLE TNF receptor-associated intracellular signaling proteins and methods of use
JOURNAL Patent: US 5563039-A 3 08-OCT-1996;
FEATURES
source location/Qualifiers
1..1384
/organism="unknown"
BASE COUNT 308 a 345 c 431 g 300 t
ORIGIN

Query Match 29.8%; Score 427.6; DB 6; Length 1384;
Best Local Similarity 78.5%; Pred. No. 4.6e-59;
Matches 540; Conservative 0; Mismatches 139; Indels 9; Gaps 2;

QY 302 GCTTCTCCGCGGCTTACCCGCGAGGGGGCGCTGCGCGCCGCTGCAAGAGCTTGGCGG 361
Db 2 GCTTCTCCAGCTTACCCGCGAGGGGGCGCTGCGCGCCGCTGCAAGAGCTTGGCGG 61
QY 363 CCGGCGTGGCCAGAGCTGGGTGCGCTGCAACTGAGACTCGGCGCGCGCCGCGAGCGGC 421
Db 62 CCGGCGTGGCCAGAGAGCGCTGCGGTGCACTGAGAGTTGGTGGAGAGTGGAGAGCG 121
QY 422 TGAAGCTTTGCTGCGGAGCAGAGAGCGCTGTTGAGTTGATCCTTACCCAGAGCCCG 481
Db 122 TGAAGATTGGCTGATGATGAGAGCGCTGTTGATTTACATCTTACCCAGAGCCCG 181
QY 482 ACCGCGCTCCGGAGTAAAGAACTGCTGAGCTGAGAGATGCGCTGGAAATCTGAAGTGC 541
Db 182 ACCGCGCTCAGGGACAGAGAACTCGCGAGCTGAGAGATGAGCTGCAAACTGAGTGTG 241
QY 542 GCTCGGGGGCCGGGATGGCCAGCGGGAGTGTGCTTGGGCCCTTGCAGCGCCGAGTGC 601
Db 242 ACTGCACTGGCCAGGGTGGAGCATACAGTATCTTCTGCAAGTTCCCGGTTT 301
QY 602 CCTCTGTGCGAGGTGAAGCGCGCGCGCCGCACTGCGCAGCTTTCTGTTCC 661
Db 302 CCTCTCCAGCGAGAGAA-----ACCACTGCGCGCGCGCGCTGCGCAGCTTTCTGTTCC 355
QY 662 AGGTGACCTGTAGTGAATGGCCGCTGAGCTTGAAGACCAACAGACGTTCCGCGCT 721
Db 356 AGGGCACTCGTATGTAACGGCCACTGACTTTCAAGACAGAGACGTTTGGCGCT 415
QY 722 CTGTGGGTCTAAATGGGCGAAGTGGGGCGCTCACTGCAACGAGGCTCGGGGCGCTGC 781
Db 416 CGGTGGGTCTCAAGTGGGCGAGGGTGGGGCGCTGCTCAACGATCACTGTGGGAGCTGA 475
QY 782 GGGAGCCGGCGCTGAGCTGCTGCGCTACAGATACGAGCGCGAGGGAGTGTACAGCAGG 841
Db 476 GAGATCTGCTTCCATCTGCTGCGCTTACAGATATGAGCTGATGGGCTATACAGCAGG 535
QY 842 CCTTCAAGCTGTGGGCGCTTGTGAGGCGCGAGGCGCGCGCCAGCGCTGACGCGC 901
Db 536 CCTTCAAGCTGTGGGCGCTTGTGAGGCGCGAGGCGCGCGCCAGCGCTGACGCGC 595
QY 902 TGGTGAAGCACTCGAGAGAAAGAGCTCACCAGCTGGC---GAGAGACTTGTGCGGCGC 958
Db 596 TGGTGAAGCGCTGAGAGAAAGAACTCACTAGTCTAGCAGAGGAGCTTGTGGGCGC 655

QY 959 TGACCATCCCAATGCGCGCTGACCTTA 986
 Db 656 AGCGGAGCCGATGCGCGCTGACCTTA 683

RESULT 12
 HSAJ11615 927 bp DNA linear PRI 16-JUL-2001
 LOCUS Homo sapiens partial TRAD gene for TNFRSF1A-associated via death
 DEFINITION domain, exons 3-4.

ACCESSION AJ311615
 VERSION AJ311615.1 GI:13940271
 KEYWORDS TNFRSF1A-associated via death domain; TRAD gene.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 927)
 AUTHORS Scheuerpflug, C.G., Dechant, M., Fellenberg, J., Ewerbeck, V. and
 Debatin, K.M.
 TITLE Sequence, genomic organisation, and mutation analysis of the human
 TRAD gene in childhood B- and T-lineage acute lymphoblastic
 leukemia and ALPS
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 927)
 AUTHORS Scheuerpflug, C.G.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAR-2001) Scheuerpflug C.G., Research, University of
 Heidelberg, Schliebacher Landstrasse 200a, 69118 Heidelberg,
 Germany, 69118, GERMANY
 COMMENT Related sequences: AJ311614-616.
 FEATURES
 location/Qualifiers
 1..927
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 gene 308..866
 /gene="TRAD"
 exon 308..585
 /gene="TRAD"
 /number=3
 /usedin=AJ311614:tradd_cds
 intron 586..667
 /gene="TRAD"
 /number=3
 668..866
 /gene="TRAD"
 /number=4
 /usedin=AJ311614:tradd_cds
 BASE COUNT 122 a 293 c 336 g 176 t
 ORIGIN

Query Match 27.8%; Score 398.4; DB 9; Length 927;
 Best Local Similarity 80.2%; Pred. No. 2,3e-54;
 Matches 518; Conservative 0; Mismatches 46; Indels 82; Gaps 1;

QY 165 GGCAGTGTACAGGAGGCTGTGAGGCTGCTGAGAGAGGCGGAGCCCGAGCGTCT 224
 Db 273 GGAACCGCCGCGGCTGTGAGGCTGTGAGGCTGTGAGAGAGGCGGAGCCCGAGCGTCT 332
 QY 225 GCAAGTCTGAAGATCCACCGAGCGAGCCCGAGAGCTGATGTCGAGCTGCGATTCTCGG 284
 Db 333 GCAGATCTGAAGATCCACCGAGCGAGCCCGAGAGCTGATGTCGAGCTGCGATTCTCGG 392
 QY 285 GCGGACGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGAGGCGGCTGTGAGGCTGTG 344
 Db 393 GCGGACGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGAGGCGGCTGTGAGGCTGTG 452
 QY 345 GCAGAGGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 404
 Db 453 GCAGAGGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 512
 QY 405 CGCGGCGCGGAGCGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 464

Db 513 CGCGGCGCGGAGCGGCTGTGAGGCTGTGAGGCTGTGAGGAGAGAGCGCTGTGAGTTCAT 572
 QY 465 CTTAGCCCGAGC----- 475
 Db 573 CTTAGCCCGAGCGGTGCGCGCGCTAGGGCTAGAGTGGGGTAGGGCGGGATCCCGCGT 632
 QY 476 -----AGCCGACCGGCTCCGGAGTGAAGAAC 502
 Db 633 TAACTCCCCGTAGACGCAAGTTCATCCCTTAGCCCGACCGGCTCCGGATGAAAGAAC 682
 QY 503 TGCTGTAGCTGTGAGAGATGCTGTGCGAAATCTGAAGTGTGAGGCTGTGAGGCTGTGAGG 562
 Db 693 TGCTGTAGCTGTGAGAGATGCTGTGCGAAATCTGAAGTGTGAGGCTGTGAGGCTGTGAGG 752
 QY 563 ACGGAGAGTGTGCTGTGCGCGCGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 622
 Db 753 ACGGAGAGTGTGCTGTGCGCGCGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 812
 QY 623 CGCGGCGCGGCGCGGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 682
 Db 813 CGCGGCGCGGCGCGGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 872
 QY 683 GCGCGCTGTGAGCTGTGAAAGCAAGACGAGCTGTGCGCGGCTGTGAGG 728
 Db 873 CGCAGGTGTGAGAGGCGGCGGTCAACGAGAGGTGTGAGTGTGAGGAGG 918

RESULT 13
 AC124713 237933 bp DNA linear HTG 23-JUN-2002
 LOCUS AC124713/c
 DEFINITION Mus musculus chromosome UNK clone RP24-116N2, WORKING DRAFT
 SEQUENCE, 7 unordered pieces.
 ACCESSION AC124713
 VERSION AC124713.2 GI:21539751
 KEYWORDS HTG, HTGS, PHASE1, HTGS, DRAFT.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 237933)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE The sequence of Mus musculus clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 237933)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 3 (bases 1 to 237933)
 McPherson, J.D. and Waterston, R.H.
 REFERENCE Direct Submission
 JOURNAL Submitted (23-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 On Jun 23, 2002 this sequence version replaced gi:21427971.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Contact: submissions@wustl.edu
 Project Information
 Center project name: M_BB0116N02
 ----- Summary Statistics -----
 Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.950319
 Consensus quality: 237142 bases at least Q40
 Consensus quality: 237539 bases at least Q30
 Consensus quality: 237879 bases at least Q20
 Insert size: 188000; agarose-fp


```

Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabali, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hayes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Huylk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlssohn, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Louie, R.,
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Mines, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, B., Nwokwenkwo, S., Oguh, M., Okunola, G.,
Oragunye, N., Oyiedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Uemami, K., Vaequez, L., Vera, Y., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 180835)
Worley, K.C.
Direct Submission
Submitted (07-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 180835)
Worley, K.C.
Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 19, 2002 this sequence version replaced gi:20467973.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Project name: GxSF
Center project name: CH230-135H12
Center clone name: CH230-135H12
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 123294 bases at least Q40
Consensus quality: 150934 bases at least Q30
Consensus quality: 135748 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 66 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preereared.
*
* 1 1314: contig of 1314 bp in length
* 1315 1414: gap of unknown length
* 1415 2575: contig of 1161 bp in length

```

```

2576 2675: gap of unknown length
2676 4270: contig of 1595 bp in length
4271 4271: gap of unknown length
4371 5486: contig of 1116 bp in length
5487 5587: gap of unknown length
5587 6854: contig of 1268 bp in length
6855 6954: gap of unknown length
6955 8187: contig of 1233 bp in length
8188 8287: gap of unknown length
8288 9815: contig of 1528 bp in length
9816 9915: gap of unknown length
9916 11334: contig of 1419 bp in length
11335 12514: gap of unknown length
12514 12514: contig of 1080 bp in length
12515 14035: gap of unknown length
14035 14035: contig of 1421 bp in length
14036 14135: gap of unknown length
14135 15352: contig of 1217 bp in length
15353 15452: gap of unknown length
15453 17088: contig of 1636 bp in length
17089 17188: gap of unknown length
17189 18530: contig of 1342 bp in length
18531 18531: gap of unknown length
18531 19810: contig of 1180 bp in length
19811 21697: gap of unknown length
21697 21797: contig of 1787 bp in length
21798 23605: contig of 1808 bp in length
23606 23705: gap of unknown length
23705 25371: gap of unknown length
25372 25471: gap of unknown length
25472 26748: contig of 1277 bp in length
26749 26848: gap of unknown length
26848 28303: contig of 1435 bp in length
28304 28403: gap of unknown length
28403 29757: contig of 1354 bp in length
29758 29857: gap of unknown length
29858 32189: contig of 2332 bp in length
32190 32289: gap of unknown length
32290 33955: contig of 1666 bp in length
33956 34055: gap of unknown length
34056 35591: contig of 1536 bp in length
35592 37651: gap of unknown length
37652 37655: contig of 1664 bp in length
37656 37755: gap of unknown length
37756 39536: gap of unknown length
39537 39636: gap of unknown length
39637 41600: contig of 1964 bp in length
41601 41700: gap of unknown length
41701 42821: contig of 1121 bp in length
42822 42921: gap of unknown length
42922 44830: contig of 1909 bp in length
44831 44930: gap of unknown length
44931 47127: contig of 2197 bp in length
47128 47227: gap of unknown length
47228 49514: contig of 2287 bp in length
49515 49614: gap of unknown length
49615 50854: contig of 1240 bp in length
50855 52701: contig of 1747 bp in length
52702 52801: gap of unknown length
52802 54279: contig of 1478 bp in length
54280 54379: gap of unknown length
54380 55721: contig of 1342 bp in length
55722 55821: gap of unknown length
55822 57052: contig of 1231 bp in length
57053 57152: gap of unknown length
57153 59073: contig of 1921 bp in length
59074 59173: gap of unknown length
59174 62044: contig of 2871 bp in length
62045 62144: gap of unknown length
62145 64479: contig of 2335 bp in length
64480 64579: gap of unknown length

```

```
* 64580 66834: contig of 2255 bp in length
* 66935 66934: gap of unknown length
* 66935 68756: contig of 1822 bp in length
* 68757 68856: gap of unknown length
* 68857 72006: contig of 3150 bp in length
* 72007 72106: gap of unknown length
* 72107 75196: contig of 3090 bp in length
* 75197 75296: gap of unknown length
* 75297 77560: contig of 2264 bp in length
* 77561 77661: gap of unknown length
* 77661 80355: contig of 2695 bp in length
* 80356 80455: gap of unknown length
* 80456 82230: contig of 1775 bp in length
* 82231 84912: gap of unknown length
* 84913 85012: contig of 2582 bp in length
* 85013 88417: gap of unknown length
* 88418 88517: contig of 3405 bp in length
* 88518 91442: gap of unknown length
* 91443 91542: contig of 2925 bp in length
* 91543 95387: gap of unknown length
* 95388 95487: contig of 3845 bp in length
* 95488 98904: gap of unknown length
* 98905 99004: contig of 3417 bp in length
* 99005 101628: gap of unknown length
* 101629 101728: contig of 2624 bp in length
* 101729 105532: gap of unknown length
* 105533 105632: contig of 3804 bp in length
* 105633 109920: gap of unknown length
* 109921 110020: contig of 4288 bp in length
* 110021 112467: gap of unknown length
* 112467 112467: contig of 2447 bp in length
```

Query Match 19.0%; Score 272; DB 24; Length 180835;

Best Local Similarity 61.3%; Pred. No. 1,7e-34; Matches 634; Conservative 0; Mismatches 170; Indels 230; Gaps 3;

```
QY 177 GGCCTGACGAGCTGCTTGTGACAGAGCGCGGAGCCCGACGTCGACAGTGTCTGA 236
DB 145385 GGCCTGACGAGCTGCTTGTGACAGAGCGCGGAGCCCGACGTCGACAGTGTCTGA 145444
QY 237 GATCCACCGCAGGACCGCGAGCTGATCGTGACGTCGCACTTGTGGGGGCGGACCCCTG 296
DB 145445 GATCCACCGCAGGACCGCGAGCTGATCGTGACGTCGCACTTGTGGGGGCGGACCCCTG 145504
QY 297 TGGCGGCTTCTCTCGCGGCTTACCGGAGGGGGGCGTGGCGGCGCGCTGACAGAGAGCCT 356
DB 145505 CGGCGGCTTCTCTCAAGACCTTACCGGAGGGGGGCGTGGCGGCGCGCTGACAGAGAGCCT 145564
QY 357 GCGCGGCGGCTGCGCCAGCACTCGGTGCGCTGCACTGAGACTGCGCGCGCGCGCA 416
DB 145565 GCGCGGCGGCTTGTGCCCAAGAGCACTGCGGTGCAACTGAGAGTTGCGTGTGTGCGGA 145624
QY 417 GCGGCTGACGCTTGTGTGCGGAGCAAGAGCGGCTTTTGAAGTTGATCCTTACCCCA--- 473
DB 145625 GCGGCTGACGCTTGTGTGCGGAGCAAGAGCGGCTTTTGAAGTTGATCCTTACCCCA 145684
QY 474 ----- 473
DB 145685 GGTGCTGTGTCTGTAGTTAGATGGAATGGTGGGGAATTTGCGGCCCAAAAGCCGACATT 145744
QY 474 -----GAGCCCGACCGGCTCGGAGTGAAGAACTGTGCTGAGTGAAG 516
DB 145745 ACTGTGACCCCTTTGGGAGCGCCGACCGGCTCAGGGAGCAAACTCGCGGAGCTGGAG 145804
QY 517 GATGCGCTGCGAAATTTGAAATCGGCTCGGGGCGCCGGGGTGGCGACGAGGAGTGTGCT 576
DB 145805 GATGAGTTCTGCAAACTGACGTGTGACTCCACTGCGCAGGAGTGAAGCACAACAGTAGCT 145864
QY 577 TCGGCGGCTTGTGAGCGCCCGGCGGCTCTGTGTGGAAGGTGAAGCGCGCGCGCGCGCG 636
DB 145865 CAGGAGGTTTGAAGTCTCTGTGTCTTCCGCGCGAGGAGAA-----GCCACTGCGCG 145918
QY 637 CCACCTGCCAGACTTTTCTGTTCCAGGGTCAAG----- 670
```

```
DB 145919 GCGCGCGGCGAGACTTTTCTGTCCATGTGCTCACTAGGTGAGTTTCAACAGCAGAA 145978
QY 671 ----- 670
DB 145979 CCGGGTACCAAGCAGTATGATCTAGCCCGAATAGGGCTGGGGCCAGGTATCCCGAGCG 146038
QY 671 -----CT 672
DB 146039 GGTGCGGAGTGTGGAATCTCGGGGAAACGATGCTTACTTACTCTTCTCTCTCTCC 146098
QY 673 GTATGAAATCGCGCGCTGAGCTTGAAGACCAACAGAGTTTGGCGGCTCTGTGGTGTTC 732
DB 146099 GCAGTGAACCGCGCACTGAATCTGCAAGACCAAGACGATTCGGCGCTCAGTGGGCTTC 146158
QY 733 AAATGGCGCAAGTGTGGGGCGCTCACTGACGAGGCTGCGGGGCGTTCGGGACCCGCG 792
DB 146159 AAGTGGCGAGGTAGGGCGCTCGCTGACGCTGTGAGCTGTCAAGCACTGAGGAACTTCC 146218
QY 793 CTGACCTGCTGACCTTACAGTACAGAGCGCGAGGACTGTACGAGCAGGCTTCCAGCTG 852
DB 146219 CTGACCTGCTGACCTTACAGTACAGAGCGCGAGGACTGTATGAGCGCTTCCAGCTG 146278
QY 853 CTGCGGCGCTTGTGTCAGAGCGGAGGCGCGCGCGCACGCTTGCACGCGCTGTGTGAAGCA 912
DB 146279 CTGCGCGCTTGTTCATCAAGACCGAGGCGCGCGTGCACACTGCAACGCTAGTGAAGCG 146338
QY 913 CTGAGGAGAACGAGCTACCAAGCGCTGAGAGGACTTGTGGGCTTACCGATCCCAAT 972
DB 146339 CTGAGGAGAACGAGCTACCTAAGTGTAGCAAGAAATCTGTAGGCTCAGGCGGAGCA 146398
QY 973 GCGGCGCTGCGCTTA 986
DB 146399 GCGGCGCAGGCTTA 146412
```

```
RESULT 15
LOCUS AR211525 200 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 3 from patent US 6399327.
ACCESSION AR211525
VERSION AR211525.1 GI:21514870
KEYWORDS
SOURCE
ORGANISM unknown.
REFERENCE 1 (bases 1 to 200)
AUTHORS Wallach,D., Boldin,M., Goncharov,T. and Golsteyn,Y.V.
TITLE Modulators of the function of FAS receptors and other proteins
JOURNAL Patent: US 6399327-A 3 04-JUN-2002;
FEATURES
Source location/Qualifiers
1..200
BASE COUNT 30 a 68 c 69 g 33 t
ORIGIN
```

Query Match 13.9%; Score 200; DB 6; Length 200;

Best Local Similarity 100.0%; Pred. No. 1,7e-22; Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 622 CCGCGCGCGCGCGCGCACTGCGCAGACTTTTCTGTTCAGAGGTCAAGCTGTAGTGAAT 681
DB 1 CCGCGCGCGCGCGCGCACTGCGCAGACTTTTCTGTTCAGAGGTCAAGCTGTAGTGAAT 60
QY 682 CCGCGCTGAGCTGAGAGCAACAGACGTTTGGCGGCTCTGTGGGTCTCAAAATGGGCG 741
DB 61 CCGCGCTGAGCTGAGAGCAACAGACGTTTGGCGGCTCTGTGGGTCTCAAAATGGGCG 120
QY 742 AAGGTGGGCGCTGCTGACGAGGAGGCTGCGGCGGCTTGTGGGAGCCCGCGCTGAGCTCG 801
DB 121 AAGGTGGGCGCTGCTGACGAGGAGGCTGCGGCGGCTTGTGGGAGCCCGCGCTGAGCTCG 180
QY 802 CTGGCTTACGAGTACGAGCG 821
```

Tue Feb 4 11:18:59 2003

us-09-763-748-1.rge

Page 17

Db 181 CTGGCTACGAGTACGACG 200

Search completed: February 3, 2003, 15:46:41
Job time : 4531 secs

THIS PAGE BLANK

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:21:53 ; Search time 329 Seconds
(without alignments)
9822.546 Million cell updates/sec

Title: US-09-763-748-1

Perfect score: 1435
Sequence: 1 ctggcgcgcgctgggaacca.....gataataagatatacagc 1435

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq_101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1435	100.0	1435	21	TRADD intracellular
2	1435	100.0	1441	17	Human TRADD intrac
3	1419.2	98.9	1459	22	Human protein enco
4	427.6	29.8	1384	17	Mouse TRADD intrac
5	218.4	15.2	649	24	Oligonucleotide fo
6	218.4	15.2	649	24	Oligonucleotide fo
7	198.8	13.9	649	24	Oligonucleotide fo
8	198.8	13.9	649	24	Oligonucleotide fo
9	71.2	5.0	2744	16	AAQ98470

C	10	69.2	4.8	114955	20	AAV53491	Human adenovine A1
	11	64.4	4.5	114955	20	AAV53491	Human adenovine A1
	12	63.8	4.4	799	19	AAV55831	Nucleotide sequenc
	13	63.8	4.4	1926	21	AAV50254	Epstein Barr virus
	14	63.8	4.4	1926	22	AAV82902	EBV tethering prot
	15	63.8	4.4	2880	21	AAV75454	Nucleotide sequenc
	16	63.8	4.4	2580	21	AAV64275	Epstein-Barr virus
C	17	63.8	4.4	5452	20	AAV90923	Anti-sense strand
	18	63.8	4.4	8705	20	AAV23778	Vector pshuttle DN
	19	63.8	4.4	9600	19	AAV21683	Vector plasmid PCM
	20	63.8	4.4	10380	20	AAV22248	Nucleotide sequenc
	21	63.8	4.4	10596	14	AAV51731	Plasmid pCISBON f
	22	63.8	4.4	10596	17	AAV40348	Plasmid pCISBON f
	23	63.8	4.4	10596	20	AAV15650	Nucleotide sequenc
C	24	63.8	4.4	16080	21	AAV59553	DNA clone pCEK C1.
	25	60	4.2	60	24	AAV41584	Human spliced tran
	26	59	4.1	2175	24	AAV97190	Human metalloprote
	27	57	4.0	14835	24	AAV94858	Human DNA sequence
	28	56.6	3.9	3034	24	ABV59321	Nucleotide sequenc
	29	56.2	3.9	2108	24	ABV67774	Oesophagus cancer
	30	55.6	3.9	9551	20	AAV22301	CDNA encoding a hu
	31	55.4	3.9	14800	24	ABV66291	Lung cancer relate
	32	53.4	3.7	1295	20	AAV90924	Epstein Barr Virus
	33	53.2	3.7	6288	22	AAV45090	CDNA encoding nove
	34	53	3.7	1215	19	AAV23483	Pseudomonas Xcp s
	35	53	3.7	1215	24	AAV13894	Pseudomonas alcali
	36	53	3.7	1215	24	AAV22871	Pseudomonas alcali
	37	53	3.7	17612	19	AAV23494	Pseudomonas Xpc, O
	38	53	3.7	17612	21	AAV13905	Pseudomonas alcali
	39	53	3.7	17612	22	AAV30870	Pseudomonas alcali
	40	52	3.6	17612	24	AAV22882	Pseudomonas alcali
	41	52	3.6	1156	24	AAV22051	Human apolipoprote
	42	51.6	3.6	10549	22	AAV15442	Human nervous syst
	43	51.4	3.6	1107	19	AAV75756	Human apolipopro
	44	51.4	3.6	1157	17	AAV06957	Human apolipopro
	45	51.4	3.6	1157	24	ABV95746	Gene #2244 used to

ALIGNMENTS

RESULT 1	AAZ93431	AAZ93431 standard; DNA; 1435 BP.	
XX	XX	AAZ93431;	
AC	AC	24-JUL-2000 (first entry)	
XX	XX	TRADD intracellular protein coding sequence.	
DT	DT	TRADD; TNF; tumour necrosis factor; NF-kappa-B; apoptosis;	
XX	XX	KW programmed cell death; antisense; inhibition; treatment; therapy;	
KW	KW	epitope shock; inflammation; cancer; antiinflammatory; human; ss.	
XX	XX	Homo sapiens.	
OS	OS		
XX	XX		
FH	FH	Location/Qualifiers	
FT	FT	1..987	
FT	FT	/tag= a	
XX	XX	/product= TRADD protein	
XX	XX		
PN	PN	WO200012527-A1.	
XX	XX		
PD	PD	09-MAR-2000.	
XX	XX		
PF	PF	25-AUG-1999; 99WO-US19614.	
XX	XX		
PR	PR	28-AUG-1998; 98US-0143212.	
XX	XX		
PA	PA	(ISIS-) ISIS PHARM INC.	
XX	XX		
PI	PI	Monia BP, Cowseert LM;	

XX MPI: 2000-237846/20.
DR P-PSDB: AAV83156.
XX
PT New antisense compounds that limit the expression of human TRADD
PT protein, useful in the treatment and diagnosis of cancer, inflammation
PT and septic shock
XX
PS Example 13: Page 61-64; 85pp; English.
XX
CC The intracellular protein TRADD has been identified as a critical
CC link between tumour necrosis factor (TNF) receptor binding and
CC downstream activation of NF-kappa-B. Overexpression of native TRADD
CC activates NF-kappa-B in the absence of TNF and dominant negative
CC mutants of TRADD block TNF-induced NF-kappa-B activation. A second
CC effect of TNF in many cell types is the induction of apoptosis
CC (programmed cell death). TRADD overexpression has been shown to
CC mimic TNF induction of apoptosis as well. Data indicates that TRADD
CC and other downstream effector proteins are the rate limiting step
CC of TNF action and would therefore serve as the most efficient
CC targets for inhibition of TNF-induced events. Antisense
CC oligonucleotides capable of inhibiting TRADD function may therefore
CC be useful in a number of therapeutic, diagnostic and research
CC applications. Inhibiting expression of TRADD by contacting human
CC cells or tissues with the antisense compound may be used to treat a
CC disease or condition associated with TRADD expression, for example,
CC septic shock, inflammation, or cancer. TRADD antisense
CC oligonucleotides of varying inhibitory capabilities are listed in
CC GENSEQ records AA293438-Z93517.
XX
SQ Sequence 1435 BP; 245 A; 437 C; 494 G; 259 T; 0 other;
Query Match 100.0%; Score 1435; DB 21; Length 1435;
Best Local Similarity 100.0%; Pred.No. 2.2e-257;
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 541 GGCTGGGGGCGCCGGGGGTGGGAGCGGGAGGTGCTTGCGGCCCTTCAGACCCCGGTG 600
QY 601 CCTCTCTGTGCGAGGTGAAGCCGCCGCCGCCGCCCACTGCCCCAATCTTTCTGTTTC 660
DB 601 CCTCTCTGTGCGAGGTGAAGCCGCCGCCGCCGCCCACTGCCCCAATCTTTCTGTTTC 660
QY 661 CAGGGTCAGCCCTGTAGTGAATCGACCGCTGAGCCCTGAAGAGACCAACAGATGTCGGCGC 720
DB 661 CAGGGTCAGCCCTGTAGTGAATCGACCGCTGAGCCCTGAAGAGACCAACAGATGTCGGCGC 720
QY 721 TCTGTGGGTCTCAAAATGCGCAGAGTGGGGCGCTCACTGACGAGAGGTGCGGGCGCTG 780
DB 721 TCTGTGGGTCTCAAAATGCGCAGAGTGGGGCGCTCACTGACGAGAGGTGCGGGCGCTG 780
QY 781 CCGGACCCGGCGCTGAGACTCGCTGAGCTACAGATACAGCGCGAGAGGACTGTACAGACAG 840
DB 781 CCGGACCCGGCGCTGAGACTCGCTGAGCTACAGATACAGCGCGAGAGGACTGTACAGACAG 840
QY 841 GCCTTCAGCTGCTGCGGGCGCTTGTGACGCGCAGAGCGCGCGCGCCAGCGCTGACGCGC 900
DB 841 GCCTTCAGCTGCTGCGGGCGCTTGTGACGCGCAGAGCGCGCGCGCCAGCGCTGACGCGC 900
QY 901 CTGTGAGAGGCACTGAGAGAACGAGCTCACAGCTGAGAGAGACTTGTGCGGCTG 960
DB 901 CTGTGAGAGGCACTGAGAGAACGAGCTCACAGCTGAGAGAGACTTGTGCGGCTG 960
QY 961 ACCGATCCCATATGCGGCGCTGAGTACAGACAGGGGTGACGCGCTTTTGGAGAACCTGG 1020
DB 961 ACCGATCCCATATGCGGCGCTGAGTACAGACAGGGGTGACGCGCTTTTGGAGAACCTGG 1020
QY 1021 ATGGCCTTAGGGTCTCTGCGGCTATTTGCTGAACCCCTGTCATCCACGGAGACCTGTA 1080
DB 1021 ATGGCCTTAGGGTCTCTGCGGCTATTTGCTGAACCCCTGTCATCCACGGAGACCTGTA 1080
QY 1081 AACTCCACTTGGCTTATCTGTGAGACTCTGTGCGGCGAGAGTTGATGCTTCCCGAGAG 1140
DB 1081 AACTCCACTTGGCTTATCTGTGAGACTCTGTGCGGCGAGAGTTGATGCTTCCCGAGAG 1140
QY 1141 CCAGACCACTGGGGGTGATCATTTGGGGATTCTGCTCAAGGTACTTTATATGATGTGGG 1200
DB 1141 CCAGACCACTGGGGGTGATCATTTGGGGATTCTGCTCAAGGTACTTTATATGATGTGGG 1200
QY 1201 GTGGGGGGGACTTGTGTTGGAGATCAGCTCACCCTTCCCATCCAGAGGGGGCTTA 1260
DB 1201 GTGGGGGGGACTTGTGTTGGAGATCAGCTCACCCTTCCCATCCAGAGGGGGCTTA 1260
QY 1261 CAGCCAGCCCTTACAGTTTCACTCATGAAGCACTTGAATCTTGTGTCTTGAACCTTCAT 1320
DB 1261 CAGCCAGCCCTTACAGTTTCACTCATGAAGCACTTGAATCTTGTGTCTTGAACCTTCAT 1320
QY 1321 CCTGGGTGCTGCAATATCTGAGTAGTAATAACAGAAATCAATCTTGGCTCCGCCAGC 1380
DB 1321 CCTGGGTGCTGCAATATCTGAGTAGTAATAACAGAAATCAATCTTGGCTCCGCCAGC 1380
QY 1381 TCACACTCAGCGTGGGACCCGGAATGTTAAGCAATGATTAATAAGTATTAACAGG 1435
DB 1381 TCACACTCAGCGTGGGACCCGGAATGTTAAGCAATGATTAATAAGTATTAACAGG 1435

RESULT 2
AAT41462
ID AAT41462 standard; cDNA: 1441 BP.
XX
AC AAT41462;
XX
DT 16-JAN-1997 (first entry)
XX
DE Human TRADD intracellular signalling protein cDNA.
XX Tumour necrosis factor receptor-1 associated death domain protein;
XX TRADD; signal transduction; cell growth; cell differentiation;
KW apoptosis; gene therapy; ds.


```

XX OS Homo sapiens.
XX XX Key Location/Qualifiers
XX FH CDS 49..987
XX FT /tag= a
XX PN W09630404-A1.
XX PD 03-OCT-1996.
XX PF 15-MAR-1996; 96WO-US03591.
XX PR 31-MAR-1995; 95US-0414625.
XX PA (TUL- ) TULARIK INC.
XX PI Goeddel DV, Hsu H;
XX DR MPI; 1996-455277/45.
XX DR P-PSDB; AAW05528.
XX PT Tumour necrosis factor receptor 1 associated death domain protein
XX PT (TRADD) - used to treat and diagnose diseases associated with
XX PT undesirable cell growth, migration and/or differentiation
XX PS Claim 9; Page 31-33; 42pp; English.
XX CC A cDNA sequence (AAT1462) codes for human tumour necrosis factor
XX CC receptor-1 (TNF-R1) associated death domain protein (TRADD) (AAW05528),
XX CC a member of a novel family of intracellular signalling proteins that
XX CC modulate cell growth, differentiation and apoptosis. The sequence
XX CC was derived from clones isolated from a HUVEC cDNA library using a
XX CC probe obtd. from a yeast two-hybrid system that identified proteins
XX CC directly interacting with the intracellular region of TNF-R1. The
XX CC cDNA can be utilised in recombinant TRADD prodn., as a probe to
XX CC detect e.g. the presence of TRADD death domain genes, and in gene
XX CC therapy appls.
XX SQ Sequence 1441 BP; 251 A; 437 C; 494 G; 259 T; 0 other;

Query Match 100.0%; Score 1435; DB 17; Length 1441;
Best Local Similarity 100.0%; Pred. No. 2.2e-257;
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCGGGGCGTGGGAACCCAGGCCCCCGAGGGGCGCCAGAGGTGAGTGGAGCTGGG 60
DB 1 CTGCGGGGCGTGGGAACCCAGGCCCCCGAGGGGCGCCAGAGGTGAGTGGAGCTGGG 60
QY 61 CAAAATGGGACGAAAGTGGGTGGGAGCGCATACCTGTTGTGAGTCTCGCTGGAC 120
DB 61 CAAAATGGGACGAAAGTGGGTGGGAGCGCATACCTGTTGTGAGTCTCGCTGGAC 120
QY 121 AAGGTGCTCTGTGATGCTTACGCGACCCCGAGAGAAAGTGGAGTGTACAGGGCT 180
DB 121 AAGGTGCTCTGTGATGCTTACGCGACCCCGAGAGAAAGTGGAGTGTACAGGGCT 180
QY 121 AAGGTGCTCTGTGATGCTTACGCGACCCCGAGAGAAAGTGGAGTGTACAGGGCT 180
DB 121 AAGGTGCTCTGTGATGCTTACGCGACCCCGAGAGAAAGTGGAGTGTACAGGGCT 180
QY 181 CTGAGAGCTGCTTGGAGAGAGCGGGGAGCCCGGACGTGTGCTGAGATGCTAAATC 240
DB 181 CTGAGAGCTGCTTGGAGAGAGCGGGGAGCCCGGACGTGTGCTGAGATGCTAAATC 240
QY 181 CTGAGAGCTGCTTGGAGAGAGCGGGGAGCCCGGACGTGTGCTGAGATGCTAAATC 240
DB 181 CTGAGAGCTGCTTGGAGAGAGCGGGGAGCCCGGACGTGTGCTGAGATGCTAAATC 240
QY 241 CACCGGAGCGACCCGAGCTGATGCTGAGCTGAGATTCTGGCGGGGAGCCCTGTGGC 300
DB 241 CACCGGAGCGACCCGAGCTGATGCTGAGCTGAGATTCTGGCGGGGAGCCCTGTGGC 300
QY 241 CACCGGAGCGACCCGAGCTGATGCTGAGCTGAGATTCTGGCGGGGAGCCCTGTGGC 300
DB 241 CACCGGAGCGACCCGAGCTGATGCTGAGCTGAGATTCTGGCGGGGAGCCCTGTGGC 300
QY 301 CGCTTCTCCGCGCTTACCGGAGAGGGGCGCTGGCGCCGCGCTGCAAGAGAGCTGGCG 360
DB 301 CGCTTCTCCGCGCTTACCGGAGAGGGGCGCTGGCGCCGCGCTGCAAGAGAGCTGGCG 360
QY 361 GCGCGCTCGCCAGCACTCGGTGCTCGCTGCACTGAGAGCTGCGCGCGCGCGAGCGG 420
DB 361 GCGCGCTCGCCAGCACTCGGTGCTCGCTGCACTGAGAGCTGCGCGCGCGCGAGCGG 420

```

```

QY 421 CTGAGACGCTTGTGCTGGCGAGCGAGAGCGCTGTTGAGATTGATCTTACCGCAGACGCC 480
DB 421 CTGAGACGCTTGTGCTGGCGAGCGAGAGCGCTGTTGAGATTGATCTTACCGCAGACGCC 480
QY 481 GACCGGCTCCGGGATGAAAGAACTGGCTGAGCTGAGAGATGCGCTGGGAAATCTGAAGTC 540
DB 481 GACCGGCTCCGGGATGAAAGAACTGGCTGAGCTGAGAGATGCGCTGGGAAATCTGAAGTC 540
QY 541 GCGCTCGGGGCGCCGGGGTGGCGAGCGGGGAGGTGCTTGGGCCCCCTTGGAGCCCCGGTG 600
DB 541 GCGCTCGGGGCGCCGGGGTGGCGAGCGGGGAGGTGCTTGGGCCCCCTTGGAGCCCCGGTG 600
QY 601 CCCTCTCTGTGAGAGTGAAGCGCGCGCGCGCGCCGACCTGACCCAGACTTTTCTGTTC 660
DB 601 CCCTCTCTGTGAGAGTGAAGCGCGCGCGCGCGCCGACCTGACCCAGACTTTTCTGTTC 660
QY 661 CAGGTCAGCTCTGATGAAATCGGCGGCTGAGCTGAGAGCAACAGAGCTTTCGGCGC 720
DB 661 CAGGTCAGCTCTGATGAAATCGGCGGCTGAGCTGAGAGCAACAGAGCTTTCGGCGC 720
QY 721 TCTGTGGGTCTCAAAATGAGCGCAAGTGGGGCGCTCACTGAGGAGAGCTGCGGGCGCTG 780
DB 721 TCTGTGGGTCTCAAAATGAGCGCAAGTGGGGCGCTCACTGAGGAGAGCTGCGGGCGCTG 780
QY 781 CCGAGACCGGCGCTGAGACTGCTGAGCTTACAGATACAGCGCGAGGGACTGTACGAGCAG 840
DB 781 CCGAGACCGGCGCTGAGACTGCTGAGCTTACAGATACAGCGCGAGGGACTGTACGAGCAG 840
QY 841 GCGTTCCAGCTGCTGCGGGCGCTTGTGCGAGCGCGAGGGCGCGCGCCGACGCTGACGCGC 900
DB 841 GCGTTCCAGCTGCTGCGGGCGCTTGTGCGAGCGCGAGGGCGCGCGCGCCGACGCTGACGCGC 900
QY 901 CTGTGGAGGCACTCGAGAGAAAGAGCTCACAGCGCTGGCAGAGACTTGTGGGCGCTG 960
DB 901 CTGTGGAGGCACTCGAGAGAAAGAGCTCACAGCGCTGGCAGAGACTTGTGGGCGCTG 960
QY 961 ACCGATCCCAATGCGCGCTGAGACCGAGGGGTGCAAGCGCTTGTGGAGAACTGG 1020
DB 961 ACCGATCCCAATGCGCGCTGAGACCGAGGGGTGCAAGCGCTTGTGGAGAACTGG 1020
QY 1021 ATGAGCTTAAAGTTCCTTCTGCGGCTATTGCTGAACCCCTGTCCATCCAGGAGCCCTGA 1080
DB 1021 ATGAGCTTAAAGTTCCTTCTGCGGCTATTGCTGAACCCCTGTCCATCCAGGAGCCCTGA 1080
QY 1081 AACTCCACTGGGCTATCTGCTGAGACCTGCTGGGGAGAGATTGACCTTCCCGAGAG 1140
DB 1081 AACTCCACTGGGCTATCTGCTGAGACCTGCTGGGGAGAGATTGACCTTCCCGAGAG 1140
QY 1141 CCAGACCACTGGGGGTGCAATCTTGGGGATTCTGCCCTAGAGTACTTGTATGAGTGGG 1200
DB 1141 CCAGACCACTGGGGGTGCAATCTTGGGGATTCTGCCCTAGAGTACTTGTATGAGTGGG 1200
QY 1201 GTGGGGGGGACTTGTGGAGATCAAGCTCACTTCTCCATCCAGAAAGCGGGGCTTA 1260
DB 1201 GTGGGGGGGACTTGTGGAGATCAAGCTCACTTCTCCATCCAGAAAGCGGGGCTTA 1260
QY 1261 CAGCGAGCCCTTACAGTTTACATCAATGAAGCACTTGAATTTGTGTCTTGGACTTGCAT 1320
DB 1261 CAGCGAGCCCTTACAGTTTACATCAATGAAGCACTTGAATTTGTGTCTTGGACTTGCAT 1320
QY 1321 CTTGGGTGCTGAGATTAAGTGAAGTAAACAGAAATCAATCTTGGCTGCGCCGAGC 1380
DB 1321 CTTGGGTGCTGAGATTAAGTGAAGTAAACAGAAATCAATCTTGGCTGCGCCGAGC 1380
QY 1381 TCACTACTCAGCTGGGAGACCCGAGATGTTAAGCAATGAATTAAGATTAACACGG 1435
DB 1381 TCACTACTCAGCTGGGAGACCCGAGATGTTAAGCAATGAATTAAGATTAACACGG 1435

RESULT 3
AAH99549
ID AAH99549 standard; cDNA; 1459 BP.
XX

```

AC	AAH99549;
XX	
DT	16-OCT-2001 (first entry)

DE Human protein encoding cDNA sequence SEQ ID NO:384.

KM Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KM antineoplastic; antirheumatic; antriarthritis; immunosuppressive;
KM antibacterial; endocrine; cardiac; central nervous system; virulence;
KM anti-HIV; fungicide; antimutagen; cardiovascular; antineoplastic; azoemia;
KM antileishmaniasis; haemostatic; vinylnary; antulcer; osteoporotic; eczema;
KM dermatological; antiallergic; antischismatic; antidiabetic; cytosaric;
KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KM anidnapyrric; rheumatoid arthritis; septic shock; pancreatitis;
KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
KM thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KM allergic rhinitis; diabetes; multiple sclerosis; depression;
KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KM neurological disorder; ss.

OS Homo sapiens.

PN W0200153455-A2.

PD 26-JUL-2001.

PF 22-DEC-2000; 2000WO-US35017.

PR 23-DEC-1999; 99US-0471275.

PR 25-APR-2000; 2000US-0552317.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT,

DR WPI; 2001-457603/49.

XX XX

PT Isolated human polynucleotides encoding polypeptides, useful for the
 treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX
 PS Claim 1, Page 486-487, 1217pp, English.

CC AAH91916.0 AAH99904 encode the human proteins given in AAH25225 to
CC AAH25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antiinflammatory; antihemmatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; vincicide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vilyneric;
CC antulcer; osteopathic; dermatologic; antiallergic; antasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.

SQ Sequence 1459 BP; 258 A; 437 C; 498 G; 266 T; 0 other;

Query Match	98.9%	Score	1419.2	DB	22	Length	1459
Best Local Similarity	99.7%	Pred. No.	1.8e-254				
Matches 1432		Conservative	0	Mismatches	3	Indels	1
						Gaps	1

QY	1	CTGGGGGGGCTGAGGAACCCAGGGCCCCGGCCGAGCGGCCAGAGAGGTGAGATGCGACGTGGG	60
Db	1	CTGGGGGGGGGTGGGAACCCAGGCCCCCGCCGAGGGCGGACAGAGAGTGAATGGCAGCTGGGG	60
QY	61	CAAAATGGGCAAGAAAGTGGGTGGGCGAGCGATACCTGTTTGTGGAGTCTCGGTGAC	120
Db	61	CAAAATGGGCAAGAAAGTGGGTGGGCAAGCGCATACCTGTTTGTGGAGTCTCGGTGAC	120
QY	121	AAGGTGTCTTGTGGATGTCTTACGCGCACCCCACAGCAAGGTGGCACTGTACAGGGCT	180
Db	121	AAGGTGTCTTGTGGATGTCTTACGCGCACCCCACAGCAAGGTGGCACTGTACAGGGCT	180
QY	181	CTGCAAGCTGCTTGTGGAGAGAGCGCGGAGACCCGAGCGTCTCAGATGCTGAAGATC	240
Db	181	CTGCAAGCTGCTTGTGGAGAGAGCGCGGAGACCCGAGCGTCTCAGATGCTGAAGATC	240
QY	241	CACCGAGCGACCCGCGAGCTGATCTGTGACGTGCGATTTCTGCGGGCGCGACCTGTGTGC	300
Db	241	CACCGAGCGACCCCGAGCTGATCTGTGACGTGCGATTTCTGCGGGCGCGACCTGTGTGC	300
QY	301	CGCTTCTCTCCGGCCCTTACCGGAGGGGGGCGCTGGCGGCGCGCTGTACAGAGAGCCGTGGG	360
Db	301	CGCTTCTCTCCGGCCCTTACCGGAGGGGGGCGCTGGCGGCGCGCTGTACAGAGAGCCGTGGG	360
QY	361	GCCGCGCTCGCCACGACCTCGGTGCGGTGCAACTGG-AGCTGCGCGCCGCGCCGAGCG	419
Db	361	GCCGCGCTCGCCACGACCTCGGTGCGGTGCAACTGGATCTTGCGCGCGCGCGCGAGCG	420
QY	420	GCTGGAAGCTTTGCTGTGGGAGACAGAGACCGCTGTTTGAATTTGCAATCTTACCCAGCAGCC	479
Db	421	GCTGGAAGCTTTGCTGTGGGAGACAGAGACCGCTGTTTGAATTTGCAATCTTACCCAGCAGCC	480
QY	480	CGACCGGCTCCGGAGATGAAGAACTGGCTGAGCTGAGAGATGCGCTGCGAAATCTTAAAGTG	539
Db	481	CGACCGGCTCCGGAGATGAAGAACTGGCTGAGTGGAGAGATGGCTGCGAAATCTTAAAGTG	540
QY	540	CGGCTCGGAGGCTCCGGGGTGCGAGCGGAGAGGTGCGCTTTGCGCCCCCTTTGCAGCCCCCGGT	599
Db	541	CGGCTCGGAGGCTCCGGGGTGCGAGCGGAGAGGTGCGCTTTGCGCCCCCTTTGCAGCCCCCGGT	600
QY	600	GCCCTCTCTGTGAGAGGTGAAGCGCGCGCGCGCGCGCACTGTGCCAGACTTTCTGTT	659
Db	601	GCCCTCTCTGTGAGAGGTGAAGCGCGCGCGCGCGCGCACTGTGCCAGACTTTCTGTT	660
QY	660	CCAGGATCAGCCTGTAGTGAATCGGCGGCTGAGCGCTGAAGAACCAACAGACGTTGCGCGG	719
Db	661	CCAGGATCAGCCTGTAGTGAATCGGCGGCTGAGCGCTGAAGAACCAACAGACGTTGCGCGG	720
QY	720	CTCTGTGGGTCTCAATATGGCGAAGTGTGGGGCGCTCACTGCAGCGAAGGCTGGCGGGCGCT	779
Db	721	CTCTGTGGGTCTCAATATGGCGAAGTGTGGGGCGCTCACTGCAGCGAAGGCTGGCGGGCGCT	780
QY	780	GCGGAGACCCGGGCTGAGACTCGTGGCTTACGAGTACAGAGCGCGAGGACTGTACAGACA	839
Db	781	GCGGAGACCCGGGCTGAGACTCGTGGCTTACGAGTACAGAGCGCGAGGACTGTACAGACA	840
QY	840	GCGCTTCCAGCTGCTCGGCGCTTCTGTACAGGCGGAGGCGCGCGCGCAAGCTGTGCAGCG	899
Db	841	GCGCTTCCAGCTGCTCGGCGCTTCTGTACAGGCGGAGGCGCGCGCGCAAGCTGTGCAGCG	900
QY	900	CCTGTGAGAGGACTCGAGAGGAAGACGACTCACCAAGCTGGGAGAGGACCTTGCTGGGCGCT	959
Db	901	CCTGTGAGAGGACTCGAGAGGAAGACGACTCACCAAGCTGGGAGAGGACCTTGCTGGGCGCT	960
QY	960	GACCGATCCCAATGAGCGGCTGGCCTTAGACCAAGGGGTGAGCGACACTTTTGGAGAACCTG	1019
Db	961	GACCGATCCCAATGAGCGGCTGGCCTTAGACCAAGGGGTGAGCGACACTTTTGGAGAACCTG	1020
QY	1020	GATGGCTTTAAGGTTCTCTTCTGCGGCTATTGTGTGAACCCCTGTCCATTCACGAGGACCTTG	1079
Db	1021	GATGGCTTTAAGGTTCTCTTCTGCGGCTATTGTGTGAACCCCTGTCCATTCACGAGGACCTTG	1080

QY 1080 AAACCTCAGCTTGGCTTACTGCTGAGCCTGCTGAGGAGAGCTGATTCCTTCCCGACAGA 1139
 Db 1081 AAACCTCAGCTTGGCTTACTGCTGAGCCTGCTGAGGAGAGCTGATTCCTTCCCGACAGA 1140
 QY 1140 GCCGAGCAGCTGGGGGTCATCATTTGGGAGATTTCCTCAGGTACTTTGATGAGTGTGG 1139
 Db 1141 GCCGAGCAGCTGGGGGTCATCATTTGGGAGATTTCCTCAGGTACTTTGATGAGTGTGG 1200
 QY 1200 GGTGGGGGGGAGCTTGTGGAGATCAGCCTCAGCTTCTCCATCCGAGAGCGGGGCTT 1259
 Db 1201 GGTGGGGGGGAGCTTGTGGAGATCAGCCTCAGCTTCTCCATCCGAGAGCGGGGCTT 1260
 QY 1260 ACAGCCAGCCCTTACAGTTTCACTCATGAGACACCTGATCTTGTGTCTCTGACCTTCA 1319
 Db 1261 ACAGCCAGCCCTTACAGTTTCACTCATGAGACACCTGATCTTGTGTCTCTGACCTTCA 1320
 QY 1320 TCCCTGGGTGCTGCGAGATACCTGAGTGAAGTAAACAGGATCAATCTTGTCTGCCCCAG 1379
 Db 1321 TCCCTGGGTGCTGCGAGATACCTGAGTGAAGTAAACAGGATCAATCTTGTCTGCCCCAG 1380
 QY 1380 CTCACACTCAGCGTGGGACCCCGAATGTTAGCAATGATATAAGTATACAGG 1435
 Db 1381 CTCACACTCAGCGTGGGACCCCGAATGTTAGCAATGATATAAGTATACAGG 1436
 RESULT 4
 AAT41463
 ID AAT41463 standard; cDNA; 1384 BP.
 AC AAT41463;
 XX
 DT 16-JAN-1997 (first entry)
 XX
 DE Mouse TRAD intracellular signalling protein cDNA.
 XX
 KW Tumour necrosis factor receptor-1 associated death domain protein;
 KM TRAD; signal transduction; cell growth; cell differentiation;
 XX apoptosis; gene therapy; ds.
 OS Mus sp.
 FH Key Location/Qualifiers
 FT CDS 1..684
 FT /*tag= a
 XX
 PN MO9630404-A1.
 PD 03-OCT-1996.
 PD 15-MAR-1996; 96MO-US03591.
 PF 31-MAR-1995; 95US-0414625.
 PR (TULA-) TULARIK INC.
 PA Goeddel DV, Hsu H;
 PI WPI; 1996-455277/45.
 DR P-PDB; AAM05530.
 DR
 XX
 XX Tumour necrosis factor receptor 1 associated death domain protein
 PT (TRAD) - used to treat and diagnose diseases associated with
 PT undesirable cell growth, migration and/or differentiation
 XX
 PS Disclosure; Page 34-35; 42pp; English.
 XX
 CC A cDNA sequence (AAT41463) codes for mouse tumour necrosis factor
 CC receptor-1 (TNR-R1) associated death domain protein (TRAD) (AAM05530),
 CC a member of a novel family of intracellular signalling proteins that
 CC modulate cell growth, differentiation and apoptosis. The cDNA can
 CC be utilised in recombinant murine TRAD prodn., as a probe to
 CC detect e.g. the presence of TRAD death domain genes, and in gene
 CC therapy aplys.

XX
 SQ Sequence 1384 BP; 308 A; 345 C; 431 G; 300 T; 0 other;
 Query Match 29.8%; Score 427.6; DB 17; Length 1384;
 Best Local Similarity 78.5%; Pred. No. 1.6e-70;
 Matches 540; Conservative 0; Mismatches 139; Indels 9; Gaps 2;
 QY 302 GCTTCTCCGCGCTACCGGAGAGGGGGGCTGCGCGCGCGGCTGACAGAGAGCTGCGG 361
 Db 2 GCTTCTCCGCGCTACCGGAGAGGGGGGCTGCGCGCGCGGCTGACAGAGAGCTGCGG 61
 QY 362 CCGGCTGCGCGCACTCGGTGCGCGCTGCACTGAGCTGCGCGCGCGCGCGCGG 421
 Db 62 CCGGCTGCGCGCAAGAGCGCTGCGGTGCGAGCTGAGGTTGCGGAGCGGAGCAGC 121
 QY 422 TGAAGCTTTGCTGCGCGAGAGAGCGCTGTTGAGTTGCATCTTACGCCAGAGCCG 481
 Db 122 TGAAGCTTTGCTGAGTGAAGAGCGCTGTTGAGTTGCATCTTACGCCAGAGCCG 181
 QY 482 ACCGGGCTCCGGGAGTGAAGAACTGGCTGAGAGAGTGGCTGCGGAATCTGAAGTGGC 541
 Db 182 ACCGGGCTCCGGGAGAGAGAACTGGAGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGA 241
 QY 542 GCTCGGGGGCGCGGGGTGCGAGCGGGAGGTGCTTTCGCGCCCTTTCAGACCCCGGTGC 601
 Db 242 ACTGCACTGGCCAGGGTGGAGCCATACAGTTAGCTTTCGAGGTTGCAAGTTCCCGGTTT 301
 QY 602 CTTCTCTGTGAGAGTGAAGCCCGCGCGCGCGCCCACTGCGCCAGACTTTTCTGTTC 661
 Db 302 CTTCTCTGAGCCGAGAGAA-----ACCACTGCGCGCGCGCTGCGAGACTTTTCTGTTC 355
 QY 662 AGGGTCAGCCTGAGTGAATTCGGCCGCTGAGCCTGAAGAGCAAGACGTTGCGCGGCT 721
 Db 356 AGGGTCAGCCTGAGTGAATTCGGCCGCTGAGCCTGAAGAGCAAGACGTTGCGCGGCT 415
 QY 722 CTGTGGTCTCAATATGCGGAGTGGGCGCTCACTGACAGAGGCTGCGGGCGCTGC 781
 Db 416 CGGTGGTCTCAATATGCGGAGTGGGCGCTGCTGCGAGCTGCTGCGGCACTGA 475
 QY 782 GGAACCCGCGCTGAGTCTGCTGCTGCTGAGTGAAGAGCGGAGGACTGTACGAGCAGG 841
 Db 476 GAGATCCTGCGCTGCACTGCTGCTGCTGAGTGAAGAGGAGGAGGAGGAGGAGGAGG 535
 QY 842 CTTTCAAGCTGCTGCGCGCGCTTGTGAGGCGGAGGCGCGCGCGCGCGCGCTGAGGCGC 901
 Db 536 CTTTCAAGCTGCTGCGCGCGCTTGTGAGGCGGAGGCGCGCGCGCGCGCGCTGAGGCGC 595
 QY 902 TGGTGAAGCACTGAGAGAGAGAGAGTCACTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 958
 Db 596 TGGTGAAGCGCTGAGAGAGAGAGAGTCACTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 655
 QY 959 TGAACGATCCCAATGGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 986
 Db 656 AGCGGAGCCGAGATGGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 683
 RESULT 5
 ABQ46658
 ID ABQ46658 standard; DNA; 649 BP.
 XX
 XX ABQ46658;
 AC
 XX
 DT 12-JUL-2002 (first entry)
 XX
 XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 33249.
 DE
 DE Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KM drug; side effect; cancer; central nervous system; cardiovascular;
 KM gastrointestinal; respiratory system; single nucleotide polymorphism;
 KM SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX

XX	MO200218632-A2.
PD	07-MAR-2002.
XX	
PF	01-SEP-2001; 2001WO-EPI0074.
PR	01-SEP-2000; 2000DE-1043826.
PR	05-SEP-2000; 2000DE-1044543.
PA	(EPIC-) EPIGENOMICS AG.
P1	Olek A, Piepenbrock C, Berlin K, Guetig D;
DR	WPL; 2002-371829/40.
XX	
PT	Determining the degree of cytosine methylation in genomic DNA, useful
PT	for diagnosis and prognosis, comprises selective hybridization of
PT	amplicons from chemically treated DNA -
XX	
P8	Claim 12; 56pp + Sequence Listing; 56pp; German.
XX	
CC	This invention describes a novel method for determining the degree of
CC	methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC	genomic sample of DNA. The sample is treated chemically to convert
CC	cytosine (C) but not methylated C, to uracil, then part of the genomic
CC	DNA that contains the target C is amplified to form a labeled amplicon.
CC	The amplicon is hybridised to two classes, each with at least one
CC	member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC	and the degree of hybridisation to both classes is determined from the
CC	label on the amplicon. From the ratio of labels hybridised to the two
CC	classes of oligomers, the degree of methylation is calculated. The method
CC	is used: (i) for diagnosis and/or prognosis of side effects of
CC	therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC	of the central nervous, cardiovascular, gastrointestinal and respiratory
CC	systems etc., particularly by detecting mutations or single nucleotide
CC	polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC	types and for investigating cell differentiation. The method allows the
CC	methylation status of many C residues to be determined simultaneously.
CC	ABQ31410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC	method for determining the degree of cytosine methylation described in
CC	the disclosure of the invention.
XX	
SQ	Sequence 649 BP; 88 A; 74 C; 249 G; 238 T; 0 other;
	Query Match 15.2%; Score 218.4; DB 24; Length 649;
	Best Local Similarity 65.0%; Pred. No. 1e-31;
	Matches 386; Conservative 0; Mismatches 126; Indels 82; Gaps 1
QY	165 GGCACTGTACAGGGCTCTCGAGGCTCGCTTGGAGAGAGCGGCGGAGCCCGAAGCTGCT 224
DB	17 GGGATGTTTCGGGGTTCGATTCTGCTTCCTTTTCGTAAGAGAGGGGGAGATTCGGAACGTGTT 76
QY	225 GCAGATGCTGAAGATCACCGCAGCGACCCGCAGCTGATCGTGCAGCTGCAATTCGCG 284
DB	77 GTAGATGTTGAAGATTATTCGACCATTCGTACTGTTGATGCTGTAAGTGCAATTTTTCG 136
QY	285 GCGGCGACCCTGTGGCCGCTTCTCCGCGCTTACCGGAGGGGGCCTGGCGCCGCGCT 344
DB	137 GCGGTAGTTTTGTGGTCGTTTTTTTGGCGTTATCCGGAGGGGGCGTTGCGCTCCGTT 196
QY	345 GCAGAGAGAGCTTGGCGGCGCGCTCGGCCAGCACTGGTGCCTGCACTGGAAGCTGG 404
DB	197 GTTAGAGAGAGTTTGGCGGTCGCTTCGTTGTTAGTATTGCGTGTCTTGTAAATTGAGATTGG 256
QY	405 CGCGCGCGCGAGCGCGCTGGAGCTTGTGCTGGGGGAGAGAGACGCGCTTGTAGTTGCAT 464
DB	257 CGTGGCGCTGAGCGGTTGAGCTTTTGTGTGGCGAGAGAACGTTGTGTGATTGAT 316
QY	465 CCTAGCCCAAG----- 474
DB	317 TTTAGTTTACTAGTCGCGTTCGGGTTAGGGTTAGATGGGGTACGGCGGAGATTCCCGGT 376
QY	475 -----OAGCCCGACCGAGCTTCGGGATGAAGAAC 502

Dd	377	TAAAGTTTTTCGTAAGGTACGTTTATTATTTTATTTAGTTCATCGTTTCGGATGAAGAAT	436
Oy	503	TGCCTGAGCTGGAGGATGCGCTGCCAATCTGAATGTCGGCTCCGGGGGCCCGGGTGCCG	562
Dd	437	TGCTTGAGTTGAGGATGCTGTCGMAATTGAAGCGCAGTTCGGGGGTTCCGGGTGGCG	496
Oy	563	ACGGGGAGGTGCGTTCCGGCCCCCTTGGCAACCCCCGGGTGCCCTCTGTGTGAGGTGAAG	622
Dd	497	ACGGGGAGGTGCGTTCCGGTTTTTTTGTGATTTTTCGATGTTTTTTTTTGTGTGAGGTAAAT	556
Oy	623	CGCCGCGCGCGCGCACCTGCCCAGACTTTTCTGTTCCAGGATGACCTGTAG	676
Dd	557	CCTCGTCGTCGTCGTTATTGTTTAGATTTTTTTTGTAGGGTTAGTTGTAG	610
 RESULT 6 ABQ46659/c			
ID	ABQ46659	standard; DNA; 649 BP.	
XX	ABQ46659;		
DT	12-JUL-2002	(first entry)	
DE	Oligonucleotide for detecting cytosine methylation SEQ ID NO 33250.		
KW	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis; drug; side effect; cancer; central nervous system; cardiovascular; gastrointestinal; respiratory system; single nucleotide polymorphism; SNP; cell differentiation; ds.		
OS	Homo sapiens.		
PN	WO200218632-A2.		
PD	07-MAR-2002.		
XX	01-SEP-2001; 2001MO-EPI0074.		
PR	01-SEP-2000; 2000DE-1043826.		
PR	05-SEP-2000; 2000DE-1044543.		
PA	(EPIC-) EPIGENOMICS AG.		
Pt	Olek A, Piepenbrock C, Berlin K, Guetig D;		
DR	WPI; 2002-371829/40.		
PT	Determining the degree of cytosine methylation in genomic DNA, useful for diagnosis and prognosis, comprises selective hybridization of amplicons from chemically treated DNA -		
PS	Claim 12; 56pp + Sequence Listing; 56pp; German.		
XX	This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-CpG-3', present in a genomic sample of DNA. The sample is treated chemically to convert cytosine (C) but not methylated C, to uracil, then part of the genomic DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes are determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. ABQ3410-ABQ54121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in		

XX (UYWY-) UNIV WYOMING.
 PA Colgin M, Lewis RV;
 PI WPI, 1995-336970/43.
 DR P-PSDB; AAR80168.
 XX Polypeptide(s) comprising repeated unit amino acid sequences, also
 PT cDNAs - derived from minor amputate spider silk proteins and used
 PT to form spider silk fibres
 XX Claim 12; Fig 1; 86pp; English.
 XX cDNA clone, pMIS1, encoding the orb web spider minor amputate silk
 CC protein Misp1 has been identified and sequenced. Repeat unit peptides
 CC of Misp1 may form part of a larger polypeptide with an amino terminus
 CC (see AAR80184-85) and a carboxy terminus (see AAR80186-89). Repeat unit
 CC peptides make up spider silk proteins (spidroins) which in turn
 CC aggregate to form the silk fibres. Spider silk fibres have high
 CC tensile strength and significant elasticity. An isolated cDNA clone of
 CC a silk protein encoding sequence is of use to produce the protein at
 CC high yields using recombinant DNA technology.
 XX Sequence 2744 BP; 486 A; 549 C; 1059 G; 608 T; 42 other;
 SQ
 Query Match 5.0%; Score 71.2; DB 16; Length 2744;
 Best Local Similarity 43.8%; Pred. No. 0.00022;
 Matches 362; Conservative 0; Mismatches 458; Indels 6; Gaps 1;
 QY 104 TGGAGTCTCGCTGAGCAAGGTGCTCTGTCGATGCTTACGCGCACCCCGAGCAAGG 163
 DB 1199 TGGAGGCTACGGTGTGTAAGTGTATACGTCGCGAGCAGAGACTGTGAGCTGCTGG 1258
 QY 164 TGGCAGTGTACAGGCTCTGTCAGGCTGCTTGGCAGAGAGCGCGGAGCCCGACCTGC 223
 DB 1259 AGCAGGAGCTGAGAGCGCTGTGTGTTACGTCGAGTGTCTGTCTGAGAGAGAGCCGC 1318
 QY 224 TGCAGATGCTGAAGATCCACCGCAGCGACCCGAGCTGATGTGACCTGTGATTTGGC 283
 DB 1319 TGCAGGTGCTGAGACAGAGACCGCTGCGGCTGACAGAGCTGAGAGCTACGGTGTCAAG 1378
 QY 284 GCGCGCAGCCCTGTGCGCGCTTCTCCGCGCTACCGCGAGGGGCGCTGCGCGCCGCGC 343
 DB 1379 TGGGTAGGTCGCGGTGACAGAGAGCTGTGCGCGCTGTCTGTGAGACAGAGAGCTGAG 1438
 QY 344 TGCAGAGAGAGCTGCGCGCGCTGCGCCAGCACTGCTGTCGCTGCACTGAGCTGC 403
 DB 1439 CGCTGTGTGTACGCTAGAGGTGCTGTGTGCTGAGAGCTGCTCAGGCGCAGAGAGC 1498
 QY 404 GCGCGCGCGCGAGAGCGCTGACGCTTGTGTGCGGAGAGAGAGCGCTTGTGAGTTGCA 463
 DB 1499 TGGAGGCTACGGTGTCAAGGTGATACGTCGCGAGCAGAGACTGTGTCTGTGAGC 1558
 QY 464 TCTTAGCCCAAGCAGCCGAGCGCTCCGGAGT-----GAAGAATGGCTGAGCTGAGG 517
 DB 1559 TGTCTGCAACAGAGAGCGCGAGCGCTGTGTGTACGTAGAGAGTCTGTGTGAGCTGG 1618
 QY 518 ATGCGCTTGCAGAAATCTGAAGTGTGCGCTGCGGGGCGCGGGGTGCGAGCGGAGGTGCTT 577
 DB 1619 TGCCTGTGTGTGAGGAGAGTGTGAGGACACCGGTGTGTGTGATATGTGTGACAAAGCGGTTA 1678
 QY 578 CGGCGCGCTTGTGAGAGCGCGCGTGTCTGTGTGAGAGGTGAACCCCGCGCGCGCGC 637
 DB 1679 TGTGTGCGAGAGAGAGCTGTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1738
 QY 638 CACCTGCCAGACTTTTCTGTTCAGAGGCTACGCTGTAGTGAATCGGCGCGCTGAGCTGA 697
 DB 1739 TTAAGGTAGAGGT 1798
 QY 698 AGGACCAAGAGCTTGT 757
 DB 1799 TGCAGGTGTGAGAGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1858

QY 758 TGCAGAGAGCTGCGCGGCGCTGCGGAGCCGCGGCTGAGTCTGCTGAGTACG 817
 DB 1859 AAGAGCTGT 1918
 QY 818 AGCGCGAGGAGCTGTACAGACAGCCCTTCCAGCTGTGTGCGCGCTTGTGTGAGCCGAGG 877
 DB 1919 AGT 1978
 QY 878 GCGCGCGCGCAGCTGT 923
 DB 1979 AGCTGAGAGCTACGT 2024
 RESULT 10
 AAX53491/C
 ID AAX53491 standard; DNA; 114955 BP.
 XX
 AC AAX53491;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE Human adenosine A1 receptor antisense oligonucleotide fragment.
 XX
 KW Antisense oligonucleotide; multiple target; antisense treatment;
 KW impaired respiration; inflammation; lung disease;
 KW pulmonary vasoconstriction; inflammation; allergic rhinitis;
 KW acute asthma; allergy; asthma; impeded respiration;
 KW respiratory distress syndrome; pain; cystic fibrosis;
 KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;
 KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
 KW colon cancer; breast cancer; lung cancer; pancreatic cancer;
 KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
 KW prostate cancer; ss.
 XX
 OS Synthetic.
 XX
 PN MO9913886-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 17-SEP-1998; 98WO-US19419.
 XX
 PR 09-JUN-1998; 98US-0093972.
 PR 17-SEP-1997; 97US-0059160.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 XX
 PI Nyce JW;
 XX
 DR WPI, 1999-229400/19.
 XX
 PT New antisense oligonucleotides used in treatment of, e.g. pulmonary
 PT vasoconstriction
 PS
 PS Disclosure; Page 37; 120pp; English.
 XX
 CC The specification describes antisense oligonucleotides (AAX52869-X55271)
 CC directed against at least 2 mRNAs selected from target genes, coding and
 CC non-coding regions of RNAs corresponding to target genes, gene
 CC initiation codons, genomic flanking regions, intron-exon borders, the
 CC 5'-end, the 3'-end and the juxta-secton between coding and non-coding
 CC regions and all segments of RNAs encoding proteins associated with one
 CC or more diseases, conditions or mixtures. The antisense oligonucleotides
 CC may be derived from sequences AAX55180-271. These multiple target
 CC oligonucleotides (specifically AAX55180-271) can be used for the
 CC antisense treatment of diseases and conditions. Typical diseases and
 CC conditions are those associated with impaired respiration and
 CC inflammation, including lung diseases, pulmonary vasoconstriction,
 CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded
 CC respiration, respiratory distress syndrome, pain, cystic fibrosis,
 CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic
 CC obstructive pulmonary disease (COPD), and cancers such as leukemias,

CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,
 CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,
 CC hepatic metastases, as well as all types of cancers which may metastasize
 CC or have metastasized to the lungs, including breast and prostate cancer.

XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match 4.8%; Score 69.2; DB 20; Length 114955;
 Best Local Similarity 31.9%; Pred. No. 0.0006;
 Matches 266; Conservative 83; Mismatches 477; Indels 9; Gaps 2;

QY 165 GGCAGTGTACAGAGCTCTGACAGCTGCTTGGACAGAGCGCGGAGCCCGACGCTT 224

DB 104951 GCG 104892

QY 225 GCAGATGCTGAAGATCCACCGCAGCGACCTGATCTGTGCACTCCGATTTCGCGG 284

DB 104891 HNNNSGGCCVGGCGNNHNNNSVGGCCVGGCGNNHNNNSVGGCCVGGCGNNHNNNS 104832

QY 285 GCGGAGAGCCCTGTGGCGCGCTTCTCTCCGCGCTTACCGAGAGGAGGAGCGTCCGCGCGCT 344

DB 104831 GCGCGCGCGNNHNNNSGCCVGGCGCGNNHNNNSGCCVGGCGCGNNHNNNSGCC 104772

QY 345 GCAGAGAGAGCTGTGGCGCGCGCGCTGCGCAGCACTCGTGCCTGCAACTGAGAGCTGCG 404

DB 104771 CVGGCGCGCGNNHNNNSGCCVGGCGCGNNHNNNSGCCVGGCGCGCGNN 104712

QY 405 CGCGCGCGCGCGCGCGCTGTGAGCGCTTGTGCGCGACAGAGAGCGCTGTTTGAAGTTGCA 464

DB 104711 HNNNSGGCG 104652

QY 465 CCTAGCCCGACAGCCCGACCGGCTCCGGATGAAAGACTGAGTGAAGATGAGATGCT 524

DB 104651 GCG 104592

QY 525 GCGAATCTGAAGTGTGCGCTCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 584

DB 104591 CCG 104532

QY 585 CTTCGAGCGCGCGCGCGCTCTCTGTGCGAGTGAAGAGCGCGCGCGCGCGCGCGCG 644

DB 104531 GCG 104472

QY 645 CCAGACTTTTCTGTTCAGAGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 704

DB 104471 GCG 104412

QY 705 ACAGAGCTTGTGCGCGCTCTGTGAGTCAAAATGCGCAAGTGG-GGCGCTCACTGACGC 763

DB 104411 CCG 104352

QY 764 GAGGCTGCGCGCGCGCTGCGCGAGACCCGCGCGCTGAGCTGCGCGCTCAAGATAC----- 816

DB 104351 GCGCGNNHNNNSGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104292

QY 817 -GAGCGCGAGGAGCTGTACAGAGAGCTTCCAGTGTGCTGCGCGCGCTTGTGAGAGCGCA 875

DB 104291 GCG 104232

QY 876 GAGCGCGCGCGCGCGCTGAGCGCGCTGTGAGAGCACTGAGAGAGAGAGCTCAACCA 935

DB 104231 HNNNSGCCCG 104172

QY 936 CCGCGAGAGAGCTTGTGAGCTGTGAGCGAGTCCCAATGCGCGCGCTTGTGAGAGC 990

DB 104171 CCG 104117

RESULT 11
 ID AAX53491 standard; DNA; 114955 BP.
 XX
 AC AAX53491;

XX 05-JUL-1999 (first entry)
 DT
 XX
 DE Human adenosine A1 receptor antisense oligonucleotide fragment.

KW Antisense oligonucleotide; multiple target; antisense treatment;
 KW impaired respiration; inflammation; lung disease;
 KW pulmonary vasoconstriction; inflammation; allergic rhinitis;
 KW acute asthma; allergy; asthma; impeded respiration;
 KW respiratory distress syndrome; pain; cystic fibrosis;
 KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;
 KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
 KW colon cancer; breast cancer; lung cancer; pancreatic cancer;
 KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
 KW prostate cancer; ss.

OS Synthetic.

PN WO913886-A1.

PD 25-MAR-1999.

PF 17-SEP-1998; 98WO-US19419.

PR 09-JUN-1998; 98US-0093972.

PR 17-SEP-1997; 97US-0059160.

PA (UVEC-) UNIV EAST CAROLINA.

PI Nyce JW;

DR WPI; 1999-229400/19.

PT New antisense oligonucleotides used in treatment of, e.g. pulmonary

PT vasoconstriction

PS Disclosure; Page 37; 120pp; English.

CC The specification describes antisense oligonucleotides (AAX52869-X55271)
 CC directed against at least 2 mRNAs selected from target genes, coding and
 CC non-coding regions of RNAs corresponding to target genes, gene
 CC initiation codons, genomic flanking regions, intron-exon borders, the
 CC 5'-end, the 3'-end and the juxta-section between coding and non-coding
 CC regions and all segments of RNAs encoding proteins associated with one
 CC or more diseases, conditions or mixtures. The antisense oligonucleotides
 CC may be derived from sequences AAX55272-74. These multiple target
 CC oligonucleotides (specifically AAX55180-271) can be used for the
 CC antisense treatment of diseases and conditions. Typical diseases and
 CC conditions are those associated with impaired respiration and
 CC inflammation, including lung diseases, pulmonary vasoconstriction,
 CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded
 CC respiration, respiratory distress syndrome, pain, cystic fibrosis,
 CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic
 CC obstructive pulmonary disease (COPD), and cancers such as leukemias,
 CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,
 CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,
 CC hepatic metastases, as well as all types of cancers which may metastasize
 CC or have metastasized to the lungs, including breast and prostate cancer.

XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match 4.5%; Score 64.4; DB 20; Length 114955;
 Best Local Similarity 32.0%; Pred. No. 0.0047;
 Matches 318; Conservative 91; Mismatches 576; Indels 10; Gaps 4;

QY 18 CAGAGCCCGCGAGCGCGCAGAGGTGAGTGCAGCTTGGCGCAAAATGGCGACAGAA 77

DB 104556 CCBGGGCG 104615

QY 78 GTGGGTGGCGAGGATACCTGTTGTGAGTCTCTGTGAGCAAGTGTGCTGTGCGGA 137

DB 104616 CBGGGCG 104675

```

QY 138 TGCCTACGCGCCACCCCGAGAGAGTGACAGGCTCTGAGGCTTGCCGCTTGCC
DB 104676 CGCGCCSNNNDNNCCGCGCGCCGCGCCSNNNDNNCCGCGCGCGCGSNNND 104735
QY 198 AGAGAGGGGGGAGCGCGGACGCTGCTGAGATGCTGAAGATCCAGCCGACCGCGCA 257
DB 104736 NNC CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104795
QY 258 GCTGATCGTGCAGCTGATTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 317
DB 104796 CSNNNDNNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGSN 104855
QY 318 CGCGAGAGGGGCGCTGCGCGCGCGCGCTGACAGAGAGCTGCGCGCGCGCTGCGCCAGCA 377
DB 104856 NNDNNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104915
QY 378 CTGCGTCCGCTGCACTGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGCTGCGC 437
DB 104916 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104975
QY 438 GAGAGAGAGCGCGCTGTTGATGCTGATCCTAGCCAGCGCGCGCGCGCGCGCT---CGGGA 494
DB 104976 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105035
QY 495 TGAAGACTGCTGAGCTGAGAGATGCGCTGCGAAT---CTGAAGTGCAGCTGCGGAGG 550
DB 105036 NDNNGGCTBGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105095
QY 551 CCGGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTGT 610
DB 105096 GCGCGSNNNDNNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105155
QY 611 CGGAGGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTGT 670
DB 105156 CGGCGCGSNNNDNNBGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105215
QY 671 CTGTAATGTAATCGCGCGCGCTGAGCTGGAAGACCAAGAGCTTCTGCTGTGAGTGC 730
DB 105216 GGGCGCGSNNNDNNBGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105275
QY 731 TCAATGCGCGCAAGTGGGCGCTGCTGCAAGCGAGGCTGCGCG--GCGCTGCGGAGACC 788
DB 105276 GSNNDNNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105335
QY 789 GGGCGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848
DB 105336 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105395
QY 849 GCTGCTGCGCGCGCTTCTGCTGAGCGCGAGGCGC--GCGCGCGCAAGCTGCTGCTGCTG 907
DB 105396 GSNNDNNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105455
QY 908 AGGCACTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 967
DB 105456 GSNNDNNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105515
QY 968 CCAATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002
DB 105516 NNDNNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105550

```

RESULT 12
AAV55831
AAV55831 standard; DNA; 799 BP.

XX AAV55831;
XX
XX 18-NOV-1998 (first entry)
XX
XX Nucleotide sequence of the stabilising sequence-encoding insert.
XX
XX Fusion protein; stabilising polypeptide; proteolytic degradation;
KW resistance; half-life; autoimmune disease; inflammation; nitro drug;

```

KW Ikappab regulator protein; inflammatory bowel disease; in vivo imaging;
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
KW cancer; pathological condition; ss.
XX
OS Epstein-Barr virus.
PN
DB WO982577-A1.
XX
DB 26-MAY-1998.
XX
PF 17-NOV-1997; 97WO-IB01508.
XX
PR 25-JUN-1997; 97US-0048945.
PR 15-NOV-1996; 96US-0030986.
XX
PA (MAST/) MASTUCCI M G.
XX
PI Masucci MG;
XX
DR WPI; 1998-312463/27.
XX
PT New fusion proteins resistant to proteolytic degradation -
PT comprising a core protein with a stabilising polypeptide comprising
PT a peptide sequence containing glycine repeats
XX
PS
XX
CC This is a nucleotide sequence of the stabilising sequence-encoding
CC insert. The invention provides a method for increasing the resistance
CC of a core protein to proteolytic degradation that comprises linking or
CC inserting onto or into the core protein a stabilising polypeptide of
CC formula (Gly)(X)(Gly)(Z)n where Gly, Glyb, Glyc are 1-6
CC sequential Gly residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met,
CC Phe, Pro or Thr and n can be anything between 1-6. X, Y and Z need not
CC be identical from n repeat to n repeat. Alternatively a nucleic acid
CC encoding the stabilising polypeptide can be linked onto or inserted into
CC a nucleic acid encoding a core protein. The fusion proteins of the
CC invention are more resistant to degradation by proteases and, thus, have
CC a longer half-life than the unfused core protein. The products can be
CC used for treating autoimmune diseases, cancer and inflammation. In
CC particular, the core protein may be an Ikappab regulator protein for the
CC treatment of inflammatory bowel disease, or a nitroreductase protein
CC which can activate nitro drugs in enzyme/prodrug therapy to treat cancer
CC or other pathological conditions. The fusion proteins can also be used in
CC diagnostic methods such as in vivo imaging.
XX
SQ Sequence 799 BP; 201 A; 106 C; 479 G; 13 T; 0 other;

```

Query Match 4.4%; Score 63.8; DB 19; Length 799;
Best Local Similarity 42.8%; Pred. No. 0.0049;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

```

QY 103 GTGAGTCTCTGCTGAGCAAGTGTCTCTGTGATGCTTACGCGGACCCCGAGAGAG 162
DB 18 GTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77
QY 163 GTGAGTGTACAGAGGCTCTGAGGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 222
DB 78 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 137
QY 223 CTGCAATGCTGAAGATTCACCGCAGCAGACCGCAGCTGATCGTGCAGTCTGCTGCT 282
DB 138 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 197
QY 283 GGGCGAGAGCTGTGAGCGCGCTTCTCGCGCGCTTACCGGAGAGAGAGAGAGAGAGAG 342
DB 198 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 257
QY 343 CTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
DB 258 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 317
QY 403 CGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462

```

Db 318 GAGGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 377
Qy 463 ATCTTACCCACACACCCACCGGCTCCGGGATGAAGACTGGAGTGAAGATGG 522
Db 378 CAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 437
Qy 523 CTGCAAAATCTGAATGCGGCTCGGGGGCCCGGGGTGCGACGCGGAGGTCCCTTCGGCC 582
Db 438 CAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 497
Qy 563 CCCTTGACAGCCCCCGGTGCTCTCTGTGTGAGAGTGAAGCCGCCGCCGCCACTT 642
Db 498 CAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 557
Qy 643 GCCCAGACTTTCTGTTCAGGGGTCAAGCTGTAAGTAATCGGCCCTGAGCTGAAGAGC 702
Db 558 GAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 617
Qy 703 CAACAGACGTTGCGCGCTCTGTGGGTCTCAATGCGCAGAGGTGGGCGCTCACTGCAG 762
Db 618 CAGAGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 677
Qy 763 CGAGGCTGCGGCGCTGTGCGGACCCGCGCTGAGACTCGCTGACCTACGATGAGCGC 822
Db 678 GGGCAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAG 737
Qy 823 GAGGACCTGTACAGAGGGGCTTCCAG 849
Db 738 GAGCGCGGGGTGAGAGGACGTGGAG 764

RESULT 13
AAA50254

ID AAA50254 standard; DNA; 1926 BP.

AC AAA50254;

DT 07-NOV-2000 (first entry)

DE Epstein Barr virus nuclear antigen 1 protein (EBNA1) DNA.

KW EBV; nuclear antigen 1; EBNA1; episome; transfection; selection;
gene therapy; ds.

OS Epstein-barr virus.

PN WO20047778-A1.

PD 17-AUG-2000.

PE 11-FEB-2000; 2000WO-US03547.

PR 11-FEB-1999; 99US-0249585.

PA (PHAR-) PHARMACOEPIA INC.

PI Horlick RA, Chelsky D;

DR WPI; 2000-515062/46.

DR P-PSDB; AAY95856.

PT Stably transfecting eukaryotic cells with at least one episome for the
production of a desired protein in vitro and for gene therapy -

PS Disclosure; Fig 2; 53pp; English.

XX The present sequence is that of DNA encoding the Epstein-Barr virus
XX (EBV) nuclear antigen protein 1 (EBNA1, see AAY95856). EBNA1 is
XX utilized in a novel method for obtaining a eukaryotic cell that is
XX stably transfected with at least one episome. This method involves
XX transfecting a eukaryotic cell with: (1) a first episome comprising
XX an EBV origin of replication (oriP, see AAA50253), a gene encoding a

CC first protein whose expression results in cell death and a
CC selectable marker for eukaryotic cells; and (2) a second episome
CC comprising an EBV oriP and a gene encoding a second protein, where
CC expression of the second protein prohibits the occurrence of cell
CC death resulting from expression of the first protein to produce
CC doubly transfected cells which also express an antigen that
CC promotes retention of the episomes by the cells. The doubly
CC transfected cells are maintained under conditions in which the
CC first and second proteins and the selectable marker are expressed,
CC and the selective pressure specified by the marker is maintained.
CC Under these conditions, only cells containing both episomes live.
CC Preferably, EBNA1 is expressed from 1 of the episomes, and the
CC protein of interest from the other episome. Either or both
CC epitopes may further comprise a nucleic acid sequence encoding a
CC protein desired to be expressed in the cell (e.g. a therapeutic
CC protein), a nucleic acid encoding an RNA that is not intended to
CC be translated (e.g. a therapeutic RNA), or a DNA sequence used as
CC a tag for the cells. The method is applicable to cell culture or
CC intact organisms, for gene therapy. It allows the rapid
CC establishment of eukaryotic cells that stably and reliably express
CC a gene of interest, using a novel method of selection, and
CC maintenance of that selection without the need for exogenous
CC selection factors, such as antibiotics.

SQ Sequence 1926 BP; 487 A; 352 C; 872 G; 215 T; 0 other;

Query Match 4.4%; Score 63.8; DB 21; Length 1926;

Best Local Similarity 42.8%; Pred. No. 0.0051;

Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

Qy 103 GTGAGTCTCTGTCGACAGGTGCTCTGTGATGCTTACGCGACCCCAAGAGAG 162
Db 260 GTGGAACAGAGACAGAGACAGAGCGGGGACAGAGCGAGAGGGGACAGAGCAGAG 319
Qy 163 GTGGAGTGTACAGGCGCTCTGACGCTGCTTGACAGAGAGCGGGAGCCGGAAGTG 222
Db 320 GAGGGGACAG 379
Qy 223 CTGCAGATGCTGAAGATCCAGCAGCAGCCGACGATGATCTGTCAGCTTGCAATTGTC 282
Db 380 GAGAGGGGACAG 439
Qy 283 GGGCGGACCCCTGTGCGCTTCTCTCCGCGCTTACCGCGAGGGGCGCTGCGCGCG 342
Db 440 GGGCAGAGCAG 499
Qy 343 CTGCAGAGAGAGCTGCGCGCGCGCTGCGCGCGAGACTGCGTGCAGACTGAGCTG 402
Db 500 GAGGGGACAG 559
Qy 403 CGCGCGGCGGACGCGGCTGAGCGCTTCTGTCGAGAGAGAGAGAGAGAGAGAGAG 462
Db 560 GAGAGGGGACAG 619
Qy 463 ATCTTACCCACACACCCACCGGCTCCGGGATGAAGACTGGAGTGAAGATGG 522
Db 620 CAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 679
Qy 523 CTGCAAAATCTGAATGCGGCTCGGGGGCCCGGGGTGCGACGCGGAGGTGCTTCCGCC 582
Db 680 CAGAGGGGACAG 739
Qy 583 CCCTTGACAGCCCCCGGTGCTCTGTGCGAGGTGAAGCCCGCGCCCGCAACTT 642
Db 740 CAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 799
Qy 643 GCCCAGACTTTCTGTTCAGGGGTACGCTGTAGTAATCGGCCGCTGAGCTTAAGAGC 702
Db 800 GAGGGGACAG 859
Qy 703 CAACAGAGTTCGCGGCTCTGTGGGTCTCAATGCGCAGAGTGGGCGCTCACTGAG 762
Db 860 CAGAGCAGAGAGGGGACAGAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 919

QY 763 CGAGGCTGCCGCGCTCGGAGACCCGCGCTGGAATCGCTGACGATACGACGC 822
DB 920 GGGAGAGAGGGGCGAGGACGAGAGAGGGGCGAGAGAGGGGCGAGGAGGTG 979
QY 823 GAGGAGCTGTACGAGCAGGCTTCCAG 849
DB 980 GAGGCCGGGGTCGAGAGGACGATGAG 1006

RESULT 14
AAFB2902
ID AAFB2902 standard; DNA; 1926 BP.

XX AAFB2902;
AC AAFB2902;
XX 29-JUN-2001 (first entry)
XX
XX BBV tethering protein EBNM1 encoding DNA.
XX
XX Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
KM Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KM EBV; latency-associated nuclear antigen; LANA; EBNM1; ds.
XX
XX Epstein-barr virus.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1926
FT /*tag= a

XX MO200125484-A2.
XX 12-APR-2001.
XX
XX 29-SEP-2000; 2000MO-US26908.
XX
XX 01-OCT-1999; 99US-0410399.
XX
XX (UNMI) UNIV MICHIGAN.
XX
XX Robertson ES, Cotter MA;
XX
XX
XX WPI; 2001-281736/29.
XX P-SDB; AAB62332.
XX
XX
XX A composition for use in gene therapy comprises an expression vector
PT that includes a nucleic acid sequence encoding a nucleic acid binding
PT protein -
XX
XX
XX Disclosure; Fig 9C; 60pp; English.

XX The invention provides a composition comprising nucleic acid, histone H1
CC protein and expression vector operably encoding a protein suitable
CC for tethering the nucleic acid to the histone H1 protein, where the
CC tethering protein is LANA. The composition is useful in aiding the
CC retention of the viral DNA in the host cell. The viral vector encodes a
CC protein suitable for tethering DNA to Histone H1. Methods for screening a
CC for compounds which are agonistic or antagonistic for the tethering of
CC viral proteins to histone H1 and DNA binding sites are useful for
CC developing the method of viral transfer. The composition has applications
CC to gene therapy, including the treatment of multiple sclerosis,
CC Parkinson's disease, Huntington disease and diabetes. The present
CC sequence represents the nucleotide sequence of the Epstein-barr virus
CC (EBV) tethering protein EBNM1.
XX
XX
XX Sequence 1926 BP; 487 A; 352 C; 872 G; 215 T; 0 other;

Query Match 4.4%; Score 63.8; DB 22; Length 1926;
Best Local Similarity 42.8%; Pred. No. 0.0051;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 103 GTGAGTCTCTGTCGACCAAGGTGCTCTGTCGATGCTTAAGGCAACCCGACAGAG 162
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 260 GTGAAACAGAGCAGAGCAGAGACCGGAGGGGCGAGAGCAGAGGGGCGAGACAGAG 319
QY 163 GTGCACTGTACAGAGGGCTGTGACGGCTGCTTGGCAAGAGCGGCGGAGCCCGACGTG 222
DB 320 GAGGAGGAGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAG 379
QY 223 CTGCAATGCTGAAGATCCACCGACGACGCCCGACCTGATCTGCAAGCTTCGATTCG 282
DB 380 GAGAGGGGCGAGAGGCGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 439
QY 283 GGGCGGACCCCTGTGCGCCCTTCTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 342
DB 440 GGGCAGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGG 499
QY 343 CTGCAAGAGACCTGGGCGCGCTCGCCCTGACCACTCGGTCGCGTCAACTGAGCTG 402
DB 500 GAGGGGCGAGAGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGG 559
QY 403 CGGCGCGCGCCGAGCGGCTGAGCGCTTGGCTGGCGAGCGAGGCGCTTGGAGTTGC 462
DB 560 GAGGAGGGGCGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGG 619
QY 463 ATCTTACCTTACGACGCCCGACCGGCTCCGGATGAAGAACTGCTGAGCTGAGATGC 522
DB 620 CAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 679
QY 523 CTGCGAAATCTGAAGTGGCTGCGGGGCGCGGGGCGAGCGGGGAGGCTTGGCGCC 582
DB 680 CAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 739
QY 583 CCTTGGAGCCCGGCTGCTCTTGTGCGAGGTGAAGCCGCGCGCGCCGACCT 642
DB 740 CAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 799
QY 643 GCCCAATCTTCTTGTTCAGGCTGACCTTGAATTCGACCTGAGCTTGAAGGAC 702
DB 800 GAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 859
QY 703 CAAACAGCTTCCGCGCTGTGTGGCTCAATGGCGCAAGTGGGGCGCTCACTGCGAG 762
DB 860 CAGAGAGGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAG 919
QY 763 CGAGGCTGCCGCGCTGCGGAGACCCGCGCTGGAATCGCTGACGATACGACGC 822
DB 920 GGGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGTG 979
QY 823 GAGGAGCTGTACGAGCAGGCTTCCAG 849
DB 980 GAGGCCGGGGTCGAGAGGACGATGAG 1006

RESULT 15
AAA75454
ID AAA75454 standard; DNA; 2580 BP.

XX AAA75454;
AC AAA75454;
XX
XX 15-JUN-2001 (first entry)
XX
XX Nucleotide sequence of the Epstein Barr nuclear antigen.
DE
XX
XX Origin of replication; protein-protein interaction; replication;
KM two-hybrid system; nuclear antigen; ss.
XX
XX
XX Epstein-barr virus.
XX
XX
XX US6114111-A.
XX
XX 05-SEP-2000.
XX
XX 30-MAR-1998; 98US-0050863.
XX
XX 30-MAR-1998; 98US-0050863.

XX (RIGE-) RIGEL PHARM INC.
PA
XX Luo Y, Payan D, Huang B;
XX WPI; 2000-593546/56.
XX
XX Composition for detecting protein-protein interactions in a mammalian
PT two hybrid system comprises bait and test vector which consist of
PT selection gene, vector viral origin of replication and fusion gene -
XX
PS Disclosure; Column 17-20; 18pp; English.
XX
XX The present sequence represents the Epstein Barr nuclear antigen.
CC It is used to produce bait vectors of the invention. The specification
CC describes a compositions and methods for a genetic system of detecting
CC protein-protein interactions in a mammalian host cell. The system
CC comprises bait and test, both containing selection genes, and viral
CC origin of replications which require bound viral replication proteins
CC to effect replication. The compositions is useful for detecting an
CC interaction between a bait protein and a test protein. It is useful in
CC a mammalian two-hybrid system for detecting protein-protein interactions
CC in a mammalian host cell.
XX
SQ Sequence 2580 BP; 632 A; 512 C; 1054 G; 382 T; 0 other;

Query Match 4.4%; Score 63.8; DB 21; Length 2580;
Best Local Similarity 42.8%; Pred. No. 0.0052;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 103 GTGGAGTCTCTGCTGACAGAGGTGCTGTGGATGCTTACGCGACCCCGACGAGAAG 162
DB 643 GTGGAAACAGAGCAGAGAGAGAGCGGAGGCGCAGAGCAGAGAGGCGCAGAGAGAG 702
QY 163 GTGCACTGTATCAGGGCTCTGAGGCTGCTTGCAAGAGCGCGAGACCCGAGACGTG 222
DB 703 GAGGGGCAAGAGCAGAGAGAGGGGCAAGAGGGGCAAGAGGGGCAAGAGAGCAG 762
QY 223 CTGCAGATGCTGAGATCCACCGCAGACCCGACGCTGATGTCAGCTGCAATTCGTC 282
DB 763 GAGGAGGGGCAAGAGCAGAGAGAGGGGCAAGAGGGGCAAGAGAGGAGCAGAGAGAG 822
QY 283 GGGCGGCAAGCCCTGTGGCCGCTTCTCCGCGCTACCGGAGGGGCGCTGCGCCGCG 342
DB 823 GGGCAGAGCAGAGAGAGGGGCAAGAGGGGCAAGAGAGGCGCAGAGAGGGCGAG 882
QY 343 CTGCAGAGAGCCTGAGCGCCGCGCTCGCCAGCACTCGTGCCGCTGCACTGAGCTG 402
DB 883 GAGGGGCAAGAGCAGAGAGAGGGGCAAGAGCAGAGAGAGGGGCAAGAGAGCAG 942
QY 403 CGCGCCGCGCGCAGCGCTGAGCGCTTGTGCGGACGAGAGCGCTGTTGAGTTGC 462
DB 943 GAGGAGGGGCAAGAGAGGGGCAAGAGGGGCAAGAGAGAGAGGCGCAGAGAGGGG 1002
QY 463 ATCTAGCCCAAGACCCGAGCCGCTCCGGGATGAAGAACTGGCTGAGTGAGTGCG 522
DB 1003 CAGGAGGGGCAAGAGAGGGGCAAGAGAGGGGCAAGAGAGAGAGGGGCAAGAGGGG 1062
QY 523 CTGCGAAATCTGAATGCGGCTCGGGGGCCCGGGGTGCGACGCGGAGAGGTCCGCTT 582
DB 1063 CAGGAGGGGCAAGAGAGGGGCAAGAGGGGCAAGAGGGGCAAGAGAGGGGCAAGAG 1122
QY 583 CCTTTCAGCCCCCGGTGCTCTTGTGTGAGAGTGAAAGCCCGCGCCGCGCCACCT 642
DB 1123 CAGGAGGGGCAAGAGAGGGGCAAGAGGGGCAAGAGGGGCAAGAGAGGGGCGAG 1182
QY 643 GCCCAGACTTTTCTGTTCAGAGGTCAAGCTGTATGTAATCGCCGCTGAGCTGAAGAC 702
DB 1183 GAGGGGCAAGAGCAGAGAGAGGGGCAAGAGGGGCAAGAGCAGAGAGGGGCAAGAGGGG 1242
QY 703 CAACAGACGTTGCGCGCTCTGTGGGTCTCAATGCGCGCAAGGTGGGGCGCTCACTGCAG 762
DB 1243 CAGGAGCAGAGGGGCAAGAGGGGCAAGAGCAGAGGGGCAAGAGGGGCAAGAGCAGAG 1302

QY 763 CGAGGCTCCCGGCGCTGCGGAGCCCGGCTGAGACTCGCTTACGAGTACGAGCC 822
DB 1303 GGGCAGAGGGGCAAGAGAGGGGCAAGAGAGGGGCAAGAGAGGGGCAAGAGAGGTG 1362
QY 823 GAGGAGCTGTACGAGCAGGCGCTTCCAG 849
DB 1363 GAGGCGGGGCTCGAGAGAGCAGTGAG 1389

Search completed: February 3, 2003, 14:33:09
Job time : 463 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:23:18 / Search time 2216 Seconds
(without alignments)
10487.601 Million cell updates/sec

Title: US-09-763-748-1
Perfect score: 1435
Sequence: 1 ctgcggcgccgtgggaaccca.....gataataagatracacgg 1435

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	765	53.3	1005	10 BE563501 601334982
2	725.8	50.6	1209	13 BE795807 601590842
3	651.2	45.4	976	13 B1193533 602947477
4	633	44.1	698	9 A1399660 602947477
5	629.8	43.9	810	12 BG327643 602426638
6	620.6	43.2	627	12 BF732310 602426638

Result No.	Score	Query Length	ID	Description
1	765	53.3	1005	10 BE563501 601334982
2	725.8	50.6	1209	13 BE795807 601590842
3	651.2	45.4	976	13 B1193533 602947477
4	633	44.1	698	9 A1399660 602947477
5	629.8	43.9	810	12 BG327643 602426638
6	620.6	43.2	627	12 BF732310 602426638

ALIGNMENTS

RESULT 1
LOCUS BE563501 1005 bp mRNA EST 15-AUG-2000
DEFINITION 601334982F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:368676 5',
mRNA sequence.
ACCESSION BE563501
VERSION BE563501.1 GI:9807221
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1005)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM382 row: d column: 21
High quality sequence stop: 728.
Location/Qualifiers 1..1005


```

Db 121 TCCTGTCGATGCTTACGCGACCCCGACGAGAAAGTGCGATGACGAGGCTCTGACG 180
Qy 188 CTGCTTGGAGAGAGCGGCGGAGCGCCGAGACGCTGACAGATGCTGAGATCCACCGCA 247
Db 181 CTGCTTGGAGAGAGCGGCGGAGCGCCGAGACGCTGACAGATGCTGAGATCCACCGCA 240
Qy 248 GCGACCCGAGCTGATGTCAGCTGCGATTTCTGCGGCGCGGACGCTGTGGCGCTTCC 307
Db 241 GCGACCCGAGCTGATGTCAGCTGCGATTTCTGCGGCGCGGACGCTGTGGCGCTTCC 300
Qy 308 TCCGCGCTTACCGCGAGAGGCGGCTGCGCGCGCGGCTGCGAGAGAGCTGCGCGCGCG 367
Db 301 TCCGCGCTTACCGCGAGAGGCGGCTGCGCGCGCGGCTGCGAGAGAGCTGCGCGCGCG 360
Qy 368 TCCGCGCGACGCTGCTGCTGCGCTGCGAGCTGCGCGCGCGCGCGCGAGCGGCTGAGC 427
Db 361 TCCGCGCGACGCTGCTGCTGCGCTGCGAGCTGCGCGCGCGCGCGCGAGCGGCTGAGC 420
Qy 428 CTTTGTGCGCGAGAGAGCGCTGTTTGAAGTTGATCTTACCCAGACGCGGACCGCGC 487
Db 421 CTTTGTGCGCGAGAGAGCGCTGTTTGAAGTTGATCTTACCCAGACGCGGACCGCGC 480
Qy 488 TCCGCGATGAGAACTGCTGAGCTGAGAGATGCGCTGCCAAATCTGAATGCGCGCTCG 547
Db 481 TCCGCGATGAGAACTGCTGAGCTGAGAGATGCGCTGCCAAATCTGAATGCGCGCTCG 540
Qy 548 GGGCGCGGAGTGCGAGAGGAGGAGTCTGCGCGCGCGCGCGCGCGCGCGCGCTCTC 607
Db 541 GGGCGCGGAGTGCGAGAGGAGGAGTCTGCGCGCGCGCGCGCGCGCGCGCGCTCTC 600
Qy 608 TGTGGAAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTCT 667
Db 601 TGTGGAAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTCT 658
Qy 668 AGCTGTGATGATTCGCGCGCTGAGCGCTG -AAGACCAAGAGCTTGGCGCGCTCT-GT 725
Db 659 CAGCTGTGATGATTCGCGCGCTGAGCGCTGAAAGAGACGAGAGCTTGGCGCGCTCTGT 718
Qy 726 GGGCTCAAAATGGGCGAA--GTTGGGCGCTCACTGAGAGAGGCTGCGGCGCGCTCG 782
Db 719 GGGCTCAAAATGGGCGAAAGTGGGCGCGCTCACTGAGAGAGGCTGCGGCGCGCTCG 778
Qy 783 GGAACCGCGCTGAGCTGCTGCGCTTACGAGTACGAGCGCA-----GGACTGTACG 835
Db 779 GGAACCGCGCG -GGACTGCTGCGCTTACGAGTACGAGCGCAAGCGCAAAAGCC 837
Qy 836 AGCAGGCTTTCAGCTGCTGCGCGCTTCTGTCAGGCGGAGCGCGCGCGCGCTGCG 895
Db 838 AGGCGCCCAAGTGTGGGCGCGCAATGTCGAAAGCCCAAGGCGCGCGCGCGCTGT 897
Qy 896 AGCGCTGTGT-----GAGGCACTCGAGAGAGAGAGCTCAAGCTGCGCAAGAGACT 949
Db 898 AGCGCTGTGTGTAGGAGACCACTCGAGAGAGAGAGCTCAAGCTGCGCAAAAGACT 957
Qy 950 TGCTGGGCTGACCGATCCCAATGCGCGCTGAGCTGAGACCGAGGCTGCG 998
Db 958 GTTGGGCGCGCGAGAGACCAAGAGCGGCGCGCATTAACGCGGCGCGC 1006

```

```

RESULT 3
LOCUS B1193533 976 bp mRNA linear EST 10-JUL-2001
DEFINITION 602947477F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5090779 5',
mRNA sequence.
ACCESSION B1193533
VERSION B1193533.1 GI:14648540
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 976)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

```

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCMI853 row: 1 column: 20
High quality sequence stop: 745.
Location/Qualifiers
1. 976
/organism="Homo sapiens"
/ab_xref="taxon:9606"
/clone_1b="NIH_MGC_42"
/issue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(6). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 160 a 281 c 390 g 145 t
ORIGIN
Query Match 45.4%; Score 651.2; DB 13; Length 976;
Best Local Similarity 90.6%; Pred. No. 6.6e-127;
Matches 779; Conservative 0; Mismatches 58; Indels 23; Gaps 7;
Qy 1 CTGGCGGCGCTGGGAAACCGAGCGCCCGCGAGCGCGCAGAGAGTGAATGAGCTGGG 60
Db 9 CTGGCGGCGCTGGGAAACCGAGCGCCCGCGAGCGCGCAGAGAGTGAATGAGCTGGG 68
Qy 61 CAAATGAGGAGCAAGAGTGGTGGGAGCGCACTACTGTTGTGAGTCTTGGCTGAGC 120
Db 69 CAAATGAGGAGCAAGAGTGGTGGGAGCGCACTACTGTTGTGAGTCTTGGCTGAGC 128
Qy 121 AAGTGTCTCTGTCGATGCTTACGCGACCCCGAGCAAGAGTGGCAAGTGAAGGCT 180
Db 129 AAGTGTCTCTGTCGATGCTTACGCGACCCCGAGCAAGAGTGGCAAGTGAAGGCT 188
Qy 181 CTGAGAGCTGCTTGGAGAGAGCGGCGGAGCGCGGAGCGTGTGCGAGTGAAGATC 240
Db 189 CTGAGAGCTGCTTGGAGAGAGCGGCGGAGCGCGGAGCGTGTGCGAGTGAAGATC 248
Qy 241 CACCGAGCGACCGCGAGCTGATGTCAGCTCGATTTCTGCGGCGCGCGCTGTGCG 300
Db 249 CACCGAGCGACCGCGAGCTGATGTCAGCTCGATTTCTGCGGCGCGCGCTGTGCG 308
Qy 301 CGCTTCTCCGCGCTTACCGCGAGGCGCGCTGCGCGCGCGCTGCAAGAGAGCTTGGC 360
Db 309 CGCTTCTCCGCGCTTACCGCGAGGCGCGCTGCGCGCGCGCTGCAAGAGAGCTTGGC 368
Qy 361 GCGCGCTGCGCGAGCACTGCTGCGCTGCACTGAGAGCTGCGCGCGCGCGCGCGAGCG 420
Db 369 GCGCGCTGCGCGAGCACTGCTGCGCTGCACTGAGAGCTGCGCGCGCGCGCGCGAGCG 428
Qy 421 CTGAGCGCTTGTGTCGCGAGAGAGCGCTGTTTGAAGTTGATCTTGAAGCGAGCGCC 480
Db 429 CTGAGCGCTTGTGTCGCGAGAGAGCGCTGTTTGAAGTTGATCTTGAAGCGAGCGCC 488
Qy 481 GACCGGCTCCGGGATGAGAACTGGCTGAGCTGAGAGATGCGCTGGAAATCTGAAGTGC 540
Db 489 GACCGGCTCCGGGATGAGAACTGGCTGAGCTGAGAGATGCGCTGGAAATCTGAAGTGC 548

```

```

QY 541 GGCTCGGGGGCCCGGGGTGCGACGCGGAGAGTCTGCTTCGCGCCCTTGTGACCCCGCGTG 600
Db 549 GGCTCGGGGGCCCGGGGTGCGACGCGGAGAGTCTGCTTCGCGCCCTTGTGACCCCGCGTG 606
QY 601 CCTCTCTCTGCGAGGTGAAGCCGCGCGCGCCGACCTGCGCCAGCTTTCTGTTTC 660
Db 607 GCGCTCTCTCTGCGAGGTGAAGC-----GCCCGCGCGCGCACTGCGCAGATTCTCT 655
QY 661 CAGGGTCAGCTGTAGTGAATCGAGCGCGCTGAGCGTGAAGAACCAAGACGTTTGGCGCG 720
Db 656 CAGGTCAGCTGTAGTGAATCGAGCGCGCTGAGCGTGAAGAACCAAGACGTTTGGCGCG 711
QY 721 TCTGTGGGTCTCAAAATGCGCAGAGTGGGCGCTCACTGACAGGAGCTGCGCGCGCTG 780
Db 712 TCTGTGGGTCTCAAAATGCGCGC-AGTGGGCGCGCTCAATG---GACGAGGTGCGCGCGCG 766
QY 781 CCGGACCCGCGCGCTGAGCTGCGCTGCGCTTACGAGTACGAGCGCG-AGGAGCTGTACGAGCA 839
Db 767 CTGGGACCCGCGCGCTGAGCTGCGCTGCGCTTACGAGTACGAGCGAGGAGTGTCCGAGCA 826
QY 840 GGCCTTCAGCTGCTGCGCGC 859
Db 827 GGCCTTCAGCTGATGGCGC 846

RESULT 4
A1399660/c 598 bp mRNA linear EST 30-MAR-1999
LOCUS th2809.x1 NCI_CGAP_Prx2 Homo sapiens cDNA clone IMAGE:2119600 3'
DEFINITION similar to SW:TRAD_HUMAN Q15628 TUMOR NECROSIS FACTOR RECEPTOR TYPE
1 ASSOCIATED PROTEIN ; mRNA sequence.
A1399660
ACCESSION A1399660.1 GI:4242747
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 698)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-remail.nih.gov
Tissue procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA library preparation: M. Bento Soares, Ph.D.
cDNA library arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNI at:
www.bio.linn.gov/bbtp/image/image.html
Insert length: 864 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 458.
Location/Qualifiers
1. 698
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2119600"
/clone_lib="NCI_CGAP_Prx2"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI_CGAP_Prx2 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
985608-986759, 1101192-1101959, and 1217928-1220615)."

```

```

BASE COUNT 150 a 202 c 200 g 146 t
ORIGIN
Query Match 44.1%; Score 633; DB 9; Length 698;
Best Local Similarity 96.9%; Pred. No. 4.1e-123;
Matches 656; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 760 CAGCAGAGCTGCGCGCGCGCTGCGGAGCCGCGCTGAGCACTGCTGCGCTTACGAGTACAG 819
Db 698 CAGCAGAGCTTCCGCTGCGCTTCCGAGAGCCGCGCTGAGCACTTGTGCTTACGAGTACAG 639
QY 820 CCGAGGAGACTGTACGAGCAGGCC-TTCAGCTGCTGCGCGCTTGTGAGCGCGAGG 878
Db 638 CCGAGGAGACTGTACGAGCAGGCCCTTTCAGTGTGCTGCGCGCTTGTGAGCGCGAGGT 579
QY 879 CCGCGCGCGCAGCTGCGAGCGCGCTGCTGAGGAGCACTGAGAGAACGAGCTTACAGCTT 938
Db 578 CCGCGCGCGCAGCTGCGAGCGCGCTGCTGAGGAGCACTGAGAGAACGAGCTTACAGCTT 519
QY 939 GGCAGAGGACTTGTGCGCGCGCTGAGCCGATCCGATGCGCGCGCTGCTGAGCAGGCGTGC 998
Db 518 GGCAGAGGACTTGTGCGCGCTGAGCCGATCCGATGCGCGCGCTGCTGAGCAGGCGTGC 459
QY 999 AGCCAGCTTTTGGAGAACCTGAGTGCCTTACCTTCTGCGGCTATTGTGTAACCC 1058
Db 458 AGCCAGCTTTTGGAGAACCTGAGTGCCTTACCTTCTGCGGCTATTGTGTAACCC 399
QY 1059 CTGTCCATCCACGAGGACCTTGAATCTCACTTGGCTTATCTGCTGAGACCTGCTGGGCGAG 1118
Db 398 CTGTCCATCCACGAGGACCTTGAATCTCACTTGGCTTATCTGCTGAGACCTGCTGGGCGAG 339
QY 1119 AGTTGATTCGCTTCCCGAGGAGCCGAGCACTGCGGCGGTGCATCATTTGGGGATTCGCGCTC 1178
Db 338 AGTTGATTCGCTTCCCGAGGAGCCGAGCACTGCGGCGGTGCATCATTTGGGGATTCGCGCTC 279
QY 1179 AGGTACTTTGATAGAGTGTGGGAGGAGGAGCTTGTGAGATCAGCTTCACTTCT 1238
Db 278 AGGTACTTTGATAGAGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 219
QY 1239 CCATCCCGAGGAGCGGGCTTACAGCCAGCCCTTACAGTTTCACTGATGAGACACCTTGA 1298
Db 218 CCATCCCGAGGAGCGGGCTTACAGCCAGCCCTTACAGTTTCACTGATGAGACACCTTGA 159
QY 1299 TCTTTGGTGTCTGTGAGCTTCACTTCCGCGGTGCTGAGATACGATGAGTAAACAGGA 1358
Db 158 TCTTTGGTGTCTGTGAGCTTCACTTCCGCGGTGCTGAGATACGATGAGTAAACAGGA 99
QY 1359 ATCAATCTTGTGCTGCGCGCGCGCTGAGCACTGAGCGGTGAGCCCGAATGTAAAGATGAT 1418
Db 98 ATCAATCTTGTGCTGCGCGCGCGCGCTGAGCACTGAGCGGTGAGCCCGAATGTAAAGATGAT 39
QY 1419 AATAAAGTATTAACAGCG 1435
Db 38 AATAAAGTATTAACCGG 22

RESULT 5
BG327643 810 bp mRNA linear EST 27-FEB-2001
LOCUS 6024263F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4564736 5'
DEFINITION mRNA sequence.
ACCESSION BG327643
VERSION BG327643.1 GI:13134081
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 810)
AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

```

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Inceye Genomes, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: L1CM1280 row: k column: 09
High quality sequence stop: 622.
Location/Qualifiers

FEATURES
source
1. 810
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4564736"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 140 a 256 c 308 g 106 t

ORIGIN

Query Match 43.9%; Score 629.8; DB 12; Length 810;
Best Local Similarity 94.9%; Pred. No. 2e-122;
Matches 682; Conservative 0; Mismatches 33; Indels 4; Gaps 3;

QY 1 CTGGCGGCGTGGGAACCCAGGCCCCCGGCGGCGGAGGTTGAGTGGAGCTGGG 60
DB 14 CTGGCGGCGTGGGAACCCAGGCCCCCGGCGGCGGAGGTTGAGTGGAGCTGGG 73
QY 61 CAAATGGGACGGAAGTGGTGGGAGCGGAGTGGTGGTGGTGGTGGTGGTGGTGG 120
DB 74 CAAATGGGACGGAAGTGGTGGGAGCGGAGTGGTGGTGGTGGTGGTGGTGGTGG 133
QY 121 AAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
DB 134 AAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 193
QY 181 CTGCAAGCTGCTTGGGAGAGAGCGGGGAGCGCGGAGCTGTCGACAGTCTGAAGATC 240
DB 194 CTGCAAGCTGCTTGGGAGAGAGCGGGGAGCGCGGAGCTGTCGACAGTCTGAAGATC 253
QY 241 CACCGGAGCGACCGGAGCTGATGTCAGAGTGGATTTGGGGGCGGAGCCCTTGGGC 300
DB 254 CACCGGAGCGACCGGAGCTGATGTCAGAGTGGATTTGGGGGCGGAGCCCTTGGGC 313
QY 301 CGCTTCTCCGCGCTTACCGGAGGGGCGCTGGCGCGCGCTGACAGAGAGCTGGCG 360
DB 314 CGCTTCTCCGCGCTTACCGGAGGGGCGCTGGCGCGCGCTGACAGAGAGCTGGCG 373
QY 361 GCGCGCTCGCGCGACCTCGGTGCGCTGCACTGAGCTGCGCGCGCGGAGCGG 420
DB 374 GCGCGCTCGCGCGACCTCGGTGCGCTGCACTGAGCTGCGCGCGCGGAGCGG 433
QY 421 CTGGAACCTTGTGCTGGGAGAGAGAGCGCTGTTGAGTGGATTCAGCCCAACACCC 480
DB 434 CTGGAACCTTGTGCTGGGAGAGAGAGCGCTGTTGAGTGGATTCAGCCCAACACCC 493
QY 481 GACCGGCTCGCGAGTGAAGAACTGCTGAGCTGAGAGATGCTGCCAAATCTGAATGC 540
DB 494 GACCGGCTCGCGAGTGAAGAACTGCTGAGCTGAGAGATGCTGCCAAATCTGAATGC 553
QY 541 GCGCTGGAGGCGCGGGGAGCGAGCGGAGAGGTGCG-CTTGGCGCGCGCTTGGACGCCCGGT 599
DB 554 GCGCTGGAGGCGCGGGGAGCGAGCGGAGAGGTGCGCAACAGGAGCCCTTGGACGCCCGGT 613

QY 600 GCCCTCTGTGCGAGGTGAAGCCGCGCGCGCGCCACCTGACCTTTCTGT 659
DB 614 GCCCTCTGTGCGAGGTGAAGCCGCGCGCGCCACCTGACCTTTCTGT 673
QY 660 CAAAGGTGACCTGTA-ATGTAATCGCGCGCTGAG-CTTGAAGACCAACAGAGCTTGG 715
DB 674 CAAAGGTGACCTGTAATGTAATCGCGCGCTGAGCTTGAAGACCAACAGAGCTTGG 732

RESULT 6
BF732310
LOCUS
DEFINITION
nae09a08.x1 NCI CGAP OV18 Homo sapiens cDNA clone IMAGE:3434774.3,
similar to SW:TRAD_HUMAN 015628 TUMOR NECROSIS FACTOR RECEPTOR TYPE
1 ASSOCIATED DEATH DOMAIN PROTEIN; contains element TARI repetitive
element; mRNA sequence.
BF732310
BF732310 GI:12057385
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
TITLE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cga@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: M. Bento
Soares, Ph.D. CDNA Library Arrayed by: Christa Frange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL, send email to:
info@image.llnl.gov
Seq primer: -400p from Gibco
High quality sequence stop: 445.
Location/Qualifiers
1. 627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3434774"
/clone_lib="NCI CGAP OV18"
/tissue_type="fibrotheoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pRT73D-Pac (pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTACCAATCGAAGTGGAGCGCGCGGCGGACATTTTCTTTTCTTTT 3');
double-stranded cDNA was ligated to Eco RI adaptors
(pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pRT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 91 a 206 c 231 g 99 t

ORIGIN

Query Match 43.2%; Score 620.6; DB 12; Length 627;
Best Local Similarity 99.4%; Pred. No. 1.6e-120;
Matches 623; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 358 GCGGCGGCTCGCGCGACCTCGGTGCGCTGCACTGAGCTGCGCGCGCGGAG 417
DB 1 GCGGCGGCTCGCGCGACCTCGGTGCGCTGCACTGAGCTGCGCGCGCGGAG 60
QY 418 CGGCTGAGCCCTTGTGCTGGGAGAGAGAGCGCTGTTGAGTGGATCTTGAAGCCAGAG 477
DB 61 CGGCTGAGCCCTTGTGCTGGGAGAGAGAGCGCTGTTGAGTGGATCTTGAAGCCAGAG 120

Qy 478 CCGGACCGGCTCCGGGATGAAGAACTGAGCTGAGCTGAGAGATGCGTGGAAATCTGAAG 537
Db 121 CCGGACCGGCTCCGGGATGAAGAACTGAGCTGAGAGATGCGTGGAAATCTGAAG 180
Qy 538 TCGGCTCGGG 597
Db 181 TCGGCTCGGG 240
Qy 598 GGGGCTCTCTGTCGAGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
Db 241 GGGGCTCTCTGTCGAGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy 658 TTCCAGGGTCAACCTGTAGTGAATCGCGCGCTGAGCTTGAAGGACCAAGAGCTTGGCG 717
Db 301 TTCCAGGGTCAACCTGTAGTGAATCGCGCGCTGAGCTTGAAGGACCAAGAGCTTGGCG 360
Qy 718 CGCTCTGTGGGTCTCAAAATGGCGCAAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 777
Db 361 CGCTCTGTGGGTCTCAAAATGGCGCAAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 420
Qy 778 CTGGCGGACCGCGCGCTGAGCTCGTGGCTTACGAGTACGAGCGAGGAGCTGTACGAG 837
Db 421 CTGGCGGACCGCGCGCTGAGCTCGTGGCTTACGAGTACGAGCGAGGAGCTGTACGAG 480
Qy 838 CAGGCTTTCAGCTGCTGGCGCGCTTCTGTGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 897
Db 481 CAGGCTTTCAGCTGCTGGCGCGCTTCTGTGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 540
Qy 898 CGGCTGTGTGAGGCACTCGAGAGAAAGAGCTTCAAGCTTGGCAGAGAGACTTGTGGGG 957
Db 541 CGGCTGTGTGAGGCACTCGAGAGAAAGAGCTTCAAGCTTGGCAGAGAGACTTGTGGGG 600
Qy 958 CTGACCGATCCCAATGGCGCGCTGGCC 984
Db 601 CTGACCGATCCCAATGGCTGGCGCTGGCC 627

RESULT 7
BM973321/c 660 bp mRNA linear EST 21-MAR-2002
LOCUS BM973321
DEFINITION UT-CF-EC1-abx-e-21-0-UI s1 UT-CF-EC1 Homo sapiens cDNA clone
UT-CF-EC1-abx-e-21-0-UI 3', mRNA sequence.

ACCESSION BM973321
VERSION BM973321.1 GI:19590912
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 660)

REFERENCE 1 (bases 1 to 660)
AUTHORS Ronaldo M.F., Lennon G. and Soares M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477

COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Meleh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).

The following repetitive elements were found in this cDNA
sequence: 77-145, >LINE2
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Location/Qualifiers

source
1. 660
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-abx-e-21-0-UI"
/clone_lib="UI-CF-EC1"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pRT3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Ronaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pRT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTCTTAC.
TAG_L1B=UI-CF-EC1
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_SEQ=AAAGTCTTAC"
BASE COUNT 139 a 191 c 184 g 146 t
ORIGIN
Query Match 43.0%; Score 616.8; DB 14; Length 660;
Best Local Similarity 99.5%; Pred. No. 1e-119;
Matches 629; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 804 GGCTTACGATGACGAGCGCGGAGGAGCTTACGAGAGGCTTCCAGCTGCTCGCGGCTT 863
Db 660 GGCTTACGATGACGAGCGCGGAGGAGCTTACGAGAGGCTTCCAGCTGCTCGCGGCTT 601
Qy 864 CGTGAGGCGGAGGCG 923
Db 600 CGTGAGGCGGAGGCG 542
Qy 924 CGAGCTACAGGCTTGGAGAGAGAGCTTGGAGAGAGAGCTTGGAGAGAGAGCTTGGAG 983
Db 541 CGAGCTACAGGCTTGGAGAGAGAGAGCTTGGAGAGAGAGCTTGGAGAGAGAGCTTGGAG 482
Qy 984 CTAGACCAAGGGGTGACGAGGAGGAGCTTGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1043
Db 481 CTAGACCAAGGGGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 422
Qy 1044 GCTATTGCTGAACCCCTGTCATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1103
Db 421 GCTATTGCTGAACCCCTGTCATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362
Qy 1104 GACCTGTGGGGGAGAGTGTGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1163
Db 361 GACCTGTGGGGGAGAGTGTGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 302
Qy 1164 TGGGATTTCTGCTTACGATCTTTATGAGTGTGGGGTGGGGGAGGAGGAGGAGGAGGAG 1223
Db 301 TGGGATTTCTGCTTACGATCTTTATGAGTGTGGGGTGGGGGAGGAGGAGGAGGAGGAGGAG 242
Qy 1224 TCAGGCTACCTTCCATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1283
Db 241 TCAGGCTACCTTCCATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 182
Qy 1284 CATGAAGACCTTGAATCTTTGGTGTCTGAGCTTCACTTCCTGGTCTCAGATACCTGAG 1343
Db 181 CATGAAGACCTTGAATCTTTGGTGTCTGAGCTTCACTTCCTGGTCTCAGATACCTGAG 122
Qy 1344 TGAAGTAAACAGAGATCAATCTTGGCTTCCAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1403
Db 121 TGAAGTAAACAGAGATCAATCTTGGCTTCCAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 62

QY 1404 ATGTTAGCAATGATTAATTAACACGG 1435
 Db 61 ATGTTAGCAATGATTAATTAACACGG 30

RESULT 8
 LOCUS BG342078 788 bp mRNA linear EST 27-FEB-2001
 DEFINITION 602463023F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4575941 5',
 mRNA sequence.
 ACCESSION BG342078 GI:13148504
 VERSION BG342078.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 788)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM1287 row: n column: 06
 High quality sequence stop: 627.
 Location/Qualifiers
 1..788
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:4575941"
 /clone_lib="NIH MGC 48"
 /tissue_type="Primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

BASE COUNT 121 a 234 c 305 g 128 t
 ORIGIN

Query Match 42.8%; Score 613.8; DB 12; Length 788;
 Best Local Similarity 93.2%; Pred. No. 4.6e-119;
 Matches 709; Conservative 0; Mismatches 42; Indels 10; Gaps 6;

QY 1 CTGGCCGGCGTGGGAACCCCGAGGGGCGGAGAGTGAATGGCGCTGGG 60
 Db 14 CTGGCCGGCGTGGGAACCCCGAGGGGCGGAGAGTGAATGGCGCTGGG 72

QY 61 CAAATATGGGACGAGAGTGGGAGGCGATACCTGTTGGAGTCTCTGGAGAC 120
 Db 73 CAAATATGGGACGAGAGTGGGAGGCGATACCTGTTGGAGTCTCTGGAGAC 132

QY 121 AAGGTGTCTCTGTGATGCTTACGCGACCCCGAGAGAGTGGAGTGAACGGACT 180
 Db 133 AAGGTGTCTCTGTGATGCTTACGCGACCCCGAGAGAGTGGAGTGAACGGACT 192

QY 181 CTGAGAGCTGCTTGGAGAGAGCGGGGAGCGCGGAGCTGCTGCAATGCTAAAGATC 240
 Db 193 CTGAGAGCTGCTTGGAGAGAGCGGGGAGCGCGGAGCTGCTGCAATGCTAAAGATC 252

QY 241 CACCGAGCGACCCCGAGAGTGAATGTCAGCTGCAATTCCTGCGGGCGGACCCCTGTGAC 300
 Db 253 CACCGAGCGACCCCGAGAGTGAATGTCAGCTGCAATTCCTGCGGGCGGACCCCTGTGAC 312

QY 301 CGCTTCTCTCGCGCTTACCGCGAGGGGGCGCTGCGCGCGCTGCAAGAGACCTTGGCG 360
 Db 313 CGCTTCTCTCGCGCTTACCGCGAGGGGGCGCTGCGCGCGCTGCAAGAGACCTTGGCG 372

QY 361 GCGGCGCTGCGCGCGAGAGTGAATGTCAGCTGCAATTCCTGCGGGCGGACCCCTGTGAC 420
 Db 373 GCGGCGCTGCGCGCGAGAGTGAATGTCAGCTGCAATTCCTGCGGGCGGACCCCTGTGAC 432

QY 421 CTGAGAGCTTGTGTGGGAGAGAGAGCGCTGTTTGAATGTCATCTTACGCCAGACGCC 480
 Db 433 CTGAGAGCTTGTGTGGGAGAGAGAGCGCTGTTTGAATGTCATCTTACGCCAGACGCC 492

QY 481 GACCGGCTCCGGAGATGAAGAACTGCTGAGCTGAGAGAGTCCCTGCGAATCTGAAGTGC 540
 Db 493 GACCGGCTCCGGAGATGAAGAACTGCTGAGAGAGTCCCTGCGAATCTGAAGTGC 552

QY 541 GCGTCCGGGGGCGCGGGGGTGGCG-ACGGGAGAGTCCGCTTGGGCCCC-TTGACGCCCGCG 598
 Db 553 GCGTCCGGGGGCGCGGGGGTGGCG-ACGGGAGAGTCCGCTTGGGCCCC-TTGACGCCCGCG 612

QY 599 TGCCCTCTCTGTGAGAGTGAAGCCCGCGCGCGCGCCACCTGCGCAGACTTTCTGT 658
 Db 613 TGCCCTCTCTGTGAGAGTGAAGCCCGCGCGCGCGCGCCACCTGCGCAGACTTTCTGT 669

QY 659 TCAGAGGTCAAGCTGTGATGAATTCGCGCGCTGAGCTGAAGAACCAAGACGTTGCGGC 718
 Db 670 TCAGAGGTCAAGCTGTGATGAATTCGCGCGCTGAGCTGAAGAACCAAGACGTTGCGGC 728

QY 719 GCCTGTGGGTCTCAATGCGGAGTGGGGGCTTCACACG 759
 Db 729 -TCTGTGGGTCTCAATGCGGAGTGGGGGCTTCACACG 766

RESULT 9
 LOCUS BG746223/C 987 bp mRNA linear EST 15-MAY-2001
 DEFINITION 602723748T1 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:4850170 3',
 mRNA sequence.
 ACCESSION BG746223
 VERSION BG746223.1 GI:14056876
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 987)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Mark Watson
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM1690 row: p column: 11
 High quality sequence start: 27
 High quality sequence stop: 788.
 Location/Qualifiers
 1..987
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:4850170"
 /clone_lib="NIH MGC 113"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:

EscoRI; cDNA made by oligo-dT priming. Directionally cloned into EscoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 209 a 301 c 304 g 173 t
ORIGIN

Query Match 42.7%; Score 613.4; DB 12; Length 987;
Best Local Similarity 88.4%; Pred. No. 5.9e-119;
Matches 827; Conservative 0; Mismatches 86; Indels 22; Gaps 14;

```
Oy 488 TCCGGGATGAGAACTGCGTGTAGATGCGCTGCGAATCTGAAATGCGCGCTCG 547
Db 935 TTCCGGATGAGAACTGCGTGTAGATGCGCTGCGAATCTGAAATGCGCGCTCG 876
Oy 548 GGG-CCCGGGGTGGCGAGCGGGAGGTGCGTTCGCCCCCTTGACG-CCCGGTGCGCTC 605
Db 875 GGGTCCCGGGGTGGCGAGCGGGAGTTCGCTTGGCCCCCTTGACGCTTCCGGTCCCTC 816
Oy 606 TCTGCGGAGGAGGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT- 664
Db 815 TCTGCGGAGGAGGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT- 756
Oy 665 GTCAAGCT-GTAGTGAATCGCGCGCTGAGCTGAGGACCAAGACGTTCCGCGCGCTC- 722
Db 755 GTCAAGCTGTAGTGAATCGCGCGCGCTGAGGACCAAGACGTTCCGCGCGCTC 696
Oy 723 TGTGGGTCTCAATGCGCGAGGTGGGGCGCTCA--CTGCAAGAGGTGCGCGCGG-CT 779
Db 695 TCTGGGTCTCAATGCGCGAGGTGGGGCGCTCACTGTCAGACGAGGCGCGCGCGAGCT 636
Oy 780 GCGGGACCGCGCGCTGGAATCGTGGGCTGACAGTACAGCGCGAGGACCTGACAGCA 839
Db 635 GCGGGACCGCGCGCTGGAATCGTGGGCTGACAGTACAGCGCGAGGACCTGACAGCA 576
Oy 840 GGCC--TTCACTGCTGTCGCGCGCTTCTGCAAGCGCGAGCGCGCGCGCGCA-CTGCA 896
Db 575 GGCCATTCAGAGATGCGCGCGCGCTTCTGCAAGCGCGAGCGCGCGCGCGCACTGTCGA 516
Oy 897 GCGCGTGTGAGG----CACTCAAGAGAAAGAGCTCAAGGCTGCGCAAGAGACTTG 951
Db 515 GCGCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 456
Oy 952 CTGGGCGCTGACCGATCCGATGCGCGCGCGCTGAGACAGGGGTGACGAGCTTTTG 1011
Db 455 CTGGGCGCTGACCGATCCGATGCGCGCGCGCTGAGACAGGGGTGACGAGCTTTTG 396
Oy 1012 AGAACCTGATGAGCTTAGAGTTCCTTCTGCGG-CTATTGTGAACCCCTGTCTCATCCAC 1070
Db 395 AGAACCTGATGAGCTTAGAGTTCCTTCTGCGGTCATGAGACCCCTGTCTCATCCAC 336
Oy 1071 GGGACCTGAAAATCTCACTTGGCTTATCTGTGAGACT-CTGGGGGAGAGTTGATGCG 1129
Db 335 GGGAACTGAAAATCCCACTTGGCTTATCTGTGAGACTTGGGGGAGAGTTGATGCG 276
Oy 1130 TTCCCC--AGAGCGAGACCACTGGGGGTGATCATTTGGGGATTCTGCTAGAGTACT 1186
Db 275 CTTCCTCCATGAGCGCATGACCACTGGGGGTGATCATTTGGGGATTCTGCTAGAGTACT 216
Oy 1187 TGATAGAGTGTGGGTGGGGGGAGCTTGTGAGATGAGCTCACTTCTCCATCCC 1246
Db 215 TGATAGAGTGTGGGTGGGGGGAGCTTGTGAGATGAGCTCACTTCTCCATCCC 156
Oy 1247 AGAAGCGGGGCTTAAAGCGAGCCCTTAAAGTTCACTCATGAAAGACTTTGATCTTTG 1306
Db 155 AGAAGCGGGGCTTAAAGCGAGCCCTTAAAGTTCACTCATGAAAGACTTTGATCTTTG 96
Oy 1307 GTCTGAGATTTATCT-GGGTCTGAGATATCTGAGAGGAAGTAAAGAAATTAATC 1365
Db 95 GTCTGAGATTTATCTAGAGGTGCTGAGATATCTGAGAGGAAGTAAAGAAATTAATC 36
```

Oy 1366 TTGCTGCCCCCAGCTACACTAGCGTGGAGCC 1400
Db 35 TTGCTGCCCCCAGCTACACTAGCGTGGAGCC 1

RESULT 10
BE789709 805 bp mRNA linear EST 20-OCT-2000
LOCUS 601481483F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3884156 5',
DEFINITION mRNA sequence.
ACCESSION BE789709
VERSION BE789709.1 GI:10210907
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 805)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LMB957 row: e column: 21
High quality sequence stop: 610.
Location/Qualifiers

FEATURES

source
1..805
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3884156"
/clone_id="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI, Site 2: SalI; Cloned unidirectionally. Primer: oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."

BASE COUNT 134 a 236 c 291 g 144 t
ORIGIN

Query Match 42.5%; Score 610; DB 12; Length 805;
Best Local Similarity 91.0%; Pred. No. 2.9e-118;
Matches 728; Conservative 0; Mismatches 60; Indels 12; Gaps 7;

```
Oy 356 TGGCGCGCGCGCTGCGCGAGCACTGCGTGGCGCTGCACTGAGCTGCGCGCGCGCG 415
Db 16 TGGCGCTGGGCAAAATAGCATGCGTGGCGCTGCACTGAGCTGCGCGCGCGCGCG 75
Oy 416 AGCGGCTGAGCGCTTGTGTGGGAGAGAGCGCTGTTTGAATTGATCTTACGCCAGC 475
Db 76 AGCGGCTGAGCGCTTGTGTGGGAGAGAGCGCTGTTTGAATTGATCTTACGCCAGC 135
Oy 476 AGCCGAGCGCGCTCGGGATGAGAACTGCGTGAAGTGAAGATGCGTGGGAAATCTGA 535
Db 136 AGCCGAGCGCGCTCGGGATGAGAACTGCGTGAAGTGAAGATGCGTGGGAAATCTGA 195
Oy 536 AGTGGGCTCGGGGCGCGGGGTGGCGAGCGGGAGGTGCTTTCGCGCGCGCTTTCGAGCC 595
Db 136 AGTGGGCTCGGGGCGCGGGGTGGCGAGCGGGAGGTGCTTTCGCGCGCGCTTTCGAGCC 255
Oy 596 CGGTCCTCTCTGTGAGAGTGAAGCGCGCGCGCGCGCGCACTTCCCAAGTCTTTC 655
Db 256 CGGTCCTCTCTGTGAGAGTGAAG--GCCGCGCGCGCGCGCACTTCCCAAGTCTTTC 312
Oy 656 TGTTCAGAGGTACGCTGTAGTAATCGGCGCGCTGAGCTGAAGACCAAGACGTTTG 715
```

|||||
Db 313 TGTTCACACCGCTGTAGTATCGCGCTGAGGCTGTGAAAGACCAACAGACGTTTCG 371
Qy 716 CGGCTCTGTGGGCTCTAAATGCGCAAGTGGGGCGCTCACTGACGAGGAGCTGGCGG 775
Db 372 CGGCTCTGTGGGCTCTAAATGCGCAAGTGGGGCGCTCACTGACGAGGAGCTGGCGG 431
Qy 776 CGCTGCGGAGCCCGGCGCTGAGCTGCTGAGCTTACAGTACGAGCGCGAGGAGCTGACG 835
Db 432 CGCTGCGGAGCCCGGCGCTGAGCTGCTGAGCTTACAGTACGAGCGCGAGGAGCTGACG 490
Qy 836 AGCAGGCTTTCACAGTGTGTGGCGCTTCTGTGACGAGCGCGCGCGCGCGCGCTGC 895
Db 491 AGCAGGCTTTCACAGTGTGTGGCGCTTCTGTGACGAGCGCGCGCGCGCGCGCTGC 550
Qy 896 AGCAGCTGTGTGAGGAGCACTGAGGAGAACAGAGCTCACAGCTGCGAGGAGGAGCTTGTGCTG 955
Db 551 AGCAGCTGTGTGAGGAGCACTGAGGAGAACAGAGCTCACAGCTGCGAGGAGGAGCTTGTGCTG 609
Qy 956 GCTGACCGATCCCAATGCGCGCTGAGCTGAGCAGGAGGAGTGCAGCGAGCTTTTGAGAA 1015
Db 610 GCTGACCGATCCCAATGCGCGCTGAGCTGAGCAGGAGGAGTGCAGCGAGCTTTTGAGAA 665
Qy 1016 CTTGATGCGCTTGTGGGTTCTTCTGTGGGCTATTGTGCTGACCCCTGTTCATTCACGGGAG 1075
Db 666 CTTGATGCGCTTGTGGGTTCTTCTGTGGGCTATTGTGCTGACCCCTGTTCATTCACGGGAG 724
Qy 1076 CTTGAACTCCACTGTGGCT-ATCGCTGACGCTGCTGGGCGAGAGTTGATGCTTCC 1134
Db 725 CTTAAATTCACCTTGTGTATTGTGACCTGTGGGCGAGAAATTGATTGTTTCC 784
Qy 1135 CAGGAGCCAGACCACTGGGG 1154
Db 785 AAGGAGCCGACACCTGGGGG 804

RESULT 11
LOCUS BG717460 632 bp mRNA linear EST 08-MAY-2001
DEFINITION 602669219F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821509 5',
mRNA sequence.
ACCESSION BG717460
VERSION BG717460.1 GI:13996647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 632)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: cgaops-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM0728 row: f column: 06
High quality sequence stop: 632.
Location/Qualifiers
1. 632

FEATURES
Source
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:4821509"
/clone_1lb="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag

); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 97 a 195 c 249 g 91 t
ORIGIN

Query Match 42.2%; Score 605; DB 12; Length 632;
Best Local Similarity 100.0%; Pred. No. 3.2e-117;
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGGCGGCGGTGGGAACCCAGGCGCCGCGAGGCGGCGAGAGGTGAGTGGAGCTGGG 60
Db 28 CTGGCGGCGGTGGGAACCCAGGCGCCGCGAGGCGGCGAGAGGTGAGTGGAGCTGGG 87
Qy 61 CAAATGAGGACGAAAGAGTGGTGGGAGCGCATACCTGTTGTGAGTCTGCTGGAC 120
Db 88 CAAATGAGGACGAAAGAGTGGTGGGAGCGCATACCTGTTGTGAGTCTGCTGGAC 147
Qy 121 AAGTGTCTCTGTGGATGCTTACGCGCACTCCGAGCAAGGTGCGATGACGGCT 180
Db 148 AAGTGTCTCTGTGGATGCTTACGCGCACTCCGAGCAAGGTGCGATGACGGCT 207
Qy 181 CTGAGGCTGCTTGGGAGAGAGGCGGAGGCGGAGCGTGTGCGAGTGTGAAGATC 240
Db 208 CTGAGGCTGCTTGGGAGAGAGGCGGAGGCGGAGCGTGTGCGAGTGTGAAGATC 267
Qy 241 CACGCGACGACCCGCGAGCTGATCTGTGCAAGCTCGATTCTGCGGCGGCGAGCTGTGC 300
Db 268 CACGCGACGACCCGCGAGCTGATCTGTGCAAGCTCGATTCTGCGGCGGCGAGCTGTGC 327
Qy 301 CGCTTCTCTCGCGCTTACCGCGAGGCGGCGCTGCGCGCTGCGAGAGCTTGGC 360
Db 328 CGCTTCTCTCGCGCTTACCGCGAGGCGGCGCTGCGCGCTGCGAGAGCTTGGC 387
Qy 361 GCGCGCTGCGCCGAGCACTGCTGCGGCTGCACTGAGCTGCGGCGCGCGCGCGAGCGG 420
Db 388 GCGCGCTGCGCCGAGCACTGCTGCGGCTGCACTGAGCTGCGGCGCGCGCGCGAGCGG 447
Qy 421 CTGACGCTTGTGTGCGGAGCGAGAGCGCTGTTGAGTGTGATCTGAGCCAGAGCC 480
Db 448 CTGACGCTTGTGTGCGGAGCGAGAGCGCTGTTGAGTGTGATCTGAGCCAGAGCC 507
Qy 481 GACCGGCTCCGGATGAAGAACTGCTGAGCTGAGAGTGCCTGCGAAATCTGAAGTGC 540
Db 508 GACCGGCTCCGGATGAAGAACTGCTGAGCTGAGAGTGCCTGCGAAATCTGAAGTGC 567
Qy 541 GGTCTGGGGGCGCGGGGTGGCGAGCGGAGAGTGTGTTGCGGCCCCCTTGGAGCCCGG 600
Db 568 GGTCTGGGGGCGCGGGGTGGCGAGCGGAGAGTGTGTTGCGGCCCCCTTGGAGCCCGG 627
Qy 601 CCTTC 605
Db 628 CCTTC 632
RESULT 12
LOCUS A1144401 609 bp mRNA linear EST 26-OCT-1998
DEFINITION gb77602.x1 Soares fetal heart MbH19w Homo sapiens cDNA clone
IMAGE:1706139 3' similar to SW-TRAD HUMAN Q15628 TUMOR NECROSIS
FACTOR RECEPTOR TYPE 1 ASSOCIATED PROTEIN ; mRNA sequence.
ACCESSION A1144401
VERSION A1144401.1 GI:3666210
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 609)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1137 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amerham
 High quality sequence stop: 463.
 Location/Qualifiers
 1. 609
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1706139"
 /clone_lib="Soares_fetal_heart_NbH19W"
 /sex="Unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: heart; Vector: pT73D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGAGCGCGCCGATCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by
 M. Fatima Bonaldo. This library was constructed from the
 same fetus as the fetal lung library, Soares fetal lung
 NbH19W."
 BASE COUNT 90 a 198 c 224 g 97 t
 ORIGIN
 Query Match 42.0%; Score 602.6; DB 9; Length 609;
 Best Local Similarity 99.3%; Pred. No. 1e-116;
 Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 376 CACTGGTGGCGCTGCACTGAGCTGGCGCGCGCGCGCGCGCTTGTCTG 435
 Db 1 CACTGGTGGCGCTGCACTGAGCTGGCGCGCGCGCGCGCGCTTGTCTG 60
 Oy 436 GCGGACGAGAGCGCTGTTGATTGATCTGATCCAGCCAGCCGCGCTCCGGAT 495
 Db 61 GCGGACGAGAGCGCTGTTGATTGATCTGATCCAGCCAGCCGCGCTCCGGAT 120
 Oy 496 GAGGAACTGGCTGAGCTGAGAGATGCGCTGCCAATCTGAAGTGGCGCTCGGAGCGCCG 555
 Db 121 GAAGAACTGGCTGAGCTGAGAGATGCGCTGCCAATCTGAAGTGGCGCTCGGAGCGCCG 180
 Oy 556 GGTGGCGAGCGGAGAGTTCCTTCGCGCGCGCGCGCGCGCGCTTGTCTGCGAG 615
 Db 181 GGTGGCGAGCGGAGAGTTCCTTCGCGCGCGCGCGCGCGCGCTTGTCTGCGAG 240
 Oy 616 GTGAAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGTCTGCGAGGATGAGCTGTA 675
 Db 241 GTGAAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGTCTGCGAGGATGAGCTGTA 300
 Oy 676 GTGAATCGGCGCGCTGAGCTGAAAGCAACAGACGTTGCGCGCTCTGTGGGTCTCAAA 735
 Db 301 GTGAATCGGCGCGCTGAGCTGAAAGCAACAGACGTTGCGCGCTCTGTGGGTCTCAAA 360
 Oy 736 TGGCCCAAGTGGGGGCTCATCTGCGAGGAGGCTGCGGGCGCTGGGGAGACCGGGCGCTG 795
 Db 361 TGGCCCAAGTGGGGGCTCATCTGCGAGGAGGCTGCGGGCGCTGGGGAGACCGGGCGCTG 420
 Oy 796 GACTGCTGGCTTCAAGTACGAGCGCGAGGAGCTGTAAGAGAGGCTTTCAGAGCTGCTG 855
 Db 421 GACTGCTGGCTTCAAGTACGAGCGCGAGGAGCTGTAAGAGAGGCTTTCAGAGCTGCTG 480
 Oy 856 CGGCGCTTGTGCGAGGCGAGGAGCGCGCGCGCGCGCGCTGCAAGCGCTGTGTGAGAGGACTC 915

Db 481 CGGCGCTTGTGCGAGGCGAGGCGCGCGCGCGCGCGCGCTGCAAGCGCTGTGTGAGGACTC 540
 Oy 916 GAGGAGAACGAGCTCACACAGCGCTGGCAGAGACTTGTGGCTGACCGATCCCAATGGC 975
 Db 541 GAGGAGAACGAGCTCACACAGCGCTGGCAGAGACTTGTGGCTGACCGATCCCAATGGC 600
 Oy 976 GGCGCTGGCC 984
 Db 601 GGCGCTGGCC 609
 RESULT 13
 A1439047 609 bp mRNA linear EST 28-MAR-1999
 LOCUS c89h08.x1 NCI CGAP CL11 Homo sapiens cDNA clone IMAGE:2073375 3'
 DEFINITION similar to SW:TRAD HUMAN 015628 TUMOR NECROSIS FACTOR RECEPTOR TYPE
 1 ASSOCIATED PROTEIN; mRNA sequence.
 ACCESSION A1439047 GI:4301586
 VERSION A1439047
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 609)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
 M.D., Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www-bio.llnl.gov/bdrp/image/image.html
 Insert Length: 1151 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 416.
 Location/Qualifiers
 1. 609
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2073375"
 /clone_lib="NCI-CGAP-CL11"
 /tissue_type="B-cell, chronic lymphocytic leukemia"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGAGCGCGCGCGCGCGCGCTTGTCTTTTCTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT73 vector.
 Library is normalized, and was constructed by Bento
 Soares and M. Fatima Bonaldo."
 BASE COUNT 91 a 200 c 222 g 96 t
 ORIGIN
 Query Match 41.9%; Score 601; DB 9; Length 609;
 Best Local Similarity 99.2%; Pred. No. 2.2e-116;
 Matches 604; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Oy 376 CACTGGTGGCGCTGCACTGAGCTGGCGCGCGCGCGCGCGCTTGTCTG 435
 Db 1 CACTGGTGGCGCTGCACTGAGCTGGCGCGCGCGCGCGCGCTTGTCTG 60
 Oy 436 GCGGACGAGAGCGCTGTTGATTGATCTGATCCAGCCAGCCGCGCTCCGGAT 495
 Db 61 GCGGACGAGAGCGCTGTTGATTGATCTGATCCAGCCAGCCGCGCTCCGGAT 120

QY 496 GAAGAACTGGCTGAGCTGAGGATGCGCTGCGAATCTGAAGTGCAGGCTGCGGGCCCGG 555
 DB 121 GAAGAACTGGCTGAGCTGAGGATGCGCTGCGAATCTGAAGTGCAGGCTGCGGGCCCGG 180
 QY 556 GTGGCGAAGCGGGAGGTGCTTGGCCCTTGGAGCCCGCGTGGCTCTCTGTCGAG 615
 DB 181 GTGGCGAAGCGGGAGGTGCTTGGCCCTTGGAGCCCGCGTGGCTCTCTGTCGAG 240
 QY 616 GTGAAGCG 675
 DB 241 GTGAAGCG 300
 QY 676 GTGAATCGCGCGCTGAGCTGAGGACCAAGACAGAGCTTGGCGCGCTGTCGAGTCA 735
 DB 301 GTGAATCGCGCGCTGAGCTGAGGACCAAGACAGAGCTTGGCGCGCTGTCGAGTCA 360
 QY 736 TGGCGCAAGGTGGGGCGCTCACTGACCGAGGCTGCCGGCGCTGCGGAGCCCGCGCTG 795
 DB 361 TGGCGCAAGGTGGGGCGCTCACTGACCGAGGCTGCCGGCGCTGCGGAGCCCGCGCTG 420
 QY 796 GACTCGCTGCGCTACGAGTACGAGCGGAGGAGGAGTACGAGCGCGCTTCCAGCTCTG 855
 DB 421 GACTCGCTGCGCTACGAGTACGAGCGGAGGAGGAGTACGAGCGCGCTTCCAGCTCTG 480
 QY 856 CGGCGCTTCTGTCAGCTG 915
 DB 481 CGGCGCTTCTGTCAGCTG 540
 QY 916 GAGGAGAACGAGCTCACCGAGCTGGAGAGGAGCTTGGCGCGCTGACCGATCCCAATGCG 975
 DB 541 GAGGAGAACGAGCTCACCGAGCTGGAGAGGAGCTTGGCGCGCTGACCGATCCCAATGCG 600
 QY 976 GGCTGGCG 984
 DB 601 GGCTGGCG 609

RESULT 14
 AK018592

LOCUS AK018592 1688 bp mRNA linear HTC 19-JAN-2002

DEFINITION Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:913005N23:homolog to TUMOR NECROSIS FACTOR RECEPTOR TYPE 1 ASSOCIATED DEATH DOMAIN PROTEIN (TNFR1-ASSOCIATED DEATH DOMAIN PROTEIN), full insert sequence.

ACCESSION AK018592

VERSION AK018592.1 GI:12858377

KEYWORDS Mus musculus (strain:C57BL/6J) adult male cecum cDNA to mRNA, HTC; CAP trapper.

SOURCE Mus musculus (strain:C57BL/6J) adult male cecum cDNA library, clone:913005N23.

ORGANISM Mus musculus

REFERENCE 1
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE 20499374
 PUBMED 11042159

REFERENCE 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE 20499374
 PUBMED 11042159

REFERENCE 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE

AUTHORS

4
 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Barlow, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochane, H., Kuehl, P., Lewis, S., Matsuo, Y., Nishida, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stabily, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, U., Mazzarelli, J., Komatsu, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitlaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE

AUTHORS

5 (bases 1 to 1688)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Haneagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE

AUTHORS

Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suenro-cho, Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer '5' GAGGAGAGGATCTCGATTATTAATTAATCCCGCCCCCGCC 3'. cDNA was prepared by using triazole thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence '5' GAGGAGAGGATCTCGATTATTAATTAATCCCGCCCCCGCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmid pUC19 (+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.

Location/Qualifiers

```

source
1. 1688
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="PANTOM DB:9130005N23"
/db_xref="MGD:MGI:1904169"
/db_xref="taxon:10090"
/clone="9130005N23"
/sex="male"
/tissue type="cecum"
/clone.lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="adult"

misc_feature
1. 1688
/note="data source:SPTR, source key:Q15628, evidence:ISS
homolog to TUMOR NECROSIS FACTOR RECEPTOR TYPE 1
ASSOCIATED DEATH DOMAIN PROTEIN (TNFR1-ASSOCIATED DEATH
DOMAIN PROTEIN)"
/db_xref="MGD:MGI:1918659"

BASE COUNT      366 a      434 c      534 g      354 t
ORIGIN

Query Match      41.6%; Score 597.4; DB 11; Length 1688;
Best Local Similarity 78.2%; Pred. No. 1.5e-115;
Matches 770; Conservative 0; Mismatches 206; Indels 9; Gaps 4;

QY 3 GCGGGCGTGGAGACCCGAGCCCGGCGGCGGCGGAGGTGAGAGTGGAGCTGGGCA 62
DB 12 GCGGCGCGAGAGACCCGAGTGTGCGGCGGCGGAGGAGGAGTGGAGCTGGGCA 71
QY 63 AATGGGACGAGAGTGGTGGGAGCGGAGTGTGGTGGTGGTGGTGGTGGTGGTGG 122
DB 72 GAATGGGACGAGAGTGGTGGGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 131
QY 123 GGTGTCTGTGTGAGTGTGCTGAGCGGAGCGGAGCGGAGGAGTGTGAGGAGCTT 182
DB 132 GGTGTCTGTGTGAGTGTGAGCGGAGCGGAGCGGAGGAGTGTGAGGAGCTT 191
QY 183 GCAGGCTGTGTGGAGAGAGCGGCGGAGCGGAGCTGTGAGAGTGTGAGAGTCCA 242
DB 192 GCAGGCTGTGTGAGAGAGTGGGAGAGCTGTGAGAGTGTGAGAGTCCA 251
QY 243 CCGGAGCGGAGCGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 302
DB 252 CCGGAGCGGAGCGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 311
QY 303 CTTCCTCCGAGCTACCGGAGGAGGCGGCTGCGCGCGGCTGAGAGAGGCTGGCGG 362
DB 312 CTTCCTCCGAGCTACCGGAGGAGGCGGCTGCGCGCGGCTGAGAGAGGCTGGCGG 371
QY 363 GCGGCTGTGCGGAGCACTGGTGTGCGGCTGTGCACTGTGCGGCGGCGGAGCGCT 422
DB 372 GCGGCTGTGCGGAGCACTGGTGTGCGGCTGTGCACTGTGCGGCGGAGCGCT 430
QY 423 GGACGCTTGTGGGAGAGAGAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 482
DB 431 GGACGCTTGTGGAGAGAGAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 490
QY 483 CCGGCTCCGAGTGAAGAACTGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 542
DB 491 CCGGCTCCGAGTGAAGAACTGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 550
QY 543 CTGGGGGGCGCGGGGTGGCGAGCGGGAGGTGTGGGCGGCGGCGGCGGCGGCTG 602
DB 551 CTGGCACTGGCGAGGGGTGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 610
QY 603 CTCTCTGTGTGGAGGTGAAGCCCGCGCGCGGCGGAGCTGCGGAGAGCTTTCTG 662
DB 611 CTCTCTGTGTGGAGGTGAAGCCCGCGCGGCGGAGCTGCGGAGAGCTTTCTG 664
QY 663 GGGTCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 722
DB 665 CCGGAGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 724
QY 723 TGTGGGTCTCAATGTGGCGAGAGGTGGGGCGCTCACTGACGAGAGGCTGCGGGCGCTGG 782

```

```

DB 725 GGTGGGTCTCAATGTGGCGAGAGGTGGGGCGCTGCTCAAGTGTGAGTGTGAGTGTGAG 784
QY 783 GGACCCGCGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 842
DB 785 AGATCTGCGC-CTGCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 843
QY 843 CTT-CCAGCTGTGCGGCGCTTCTGTGAGAGCGGAGGCGCGCGGCGGAGCGTGTGAGGCGCC 901
DB 844 CTTCCTCCAGCTGTGCGGCGCTTCTGTGAGAGCGGAGGCGCGCGGCGGAGCGTGTGAGGCGCC 903
QY 902 TGTGTGAGAGCACTGTGAGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 961
DB 904 TGTGTGAGAGCGCTGTGAGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 963
QY 962 CCGATCCCAATGTGGCGGCGCTGTGAGCTTA 986
DB 964 CCGAGCGCGAGTGTGGCGGCGCTGTGAGCTTA 988

RESULT 15
BG432081      874 bp      mRNA      linear      EST 14-MAR-2001
LOCUS      60249694F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4610723 5',
DEFINITION      mRNA sequence.
ACCESSION      BG432081
VERSION      BG432081.1 GI:13338587
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 874)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1356 row: 9 column: 12
High quality sequence stop: 669.
Location/Qualifiers
1. 874
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone.lib="IMAGE:4610723"
/clone.lib="NIH MGC 75"
/lab_host="DH10B (TI phage-resistance)"
/note="Organ: Kidney; Vector: pMDR-LIB (Clontech); Site: 1;
Site 1 (ggcgctggcc); Site 2: Site 1 (ggcgctggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCCATATGAGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGAGGCGGCGGAGCATG-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."

BASE COUNT      131 a      280 c      315 g      147 t      1 others
ORIGIN

Query Match      40.9%; Score 586.2; DB 12; Length 874;
Best Local Similarity 90.8%; Pred. No. 3e-113;
Matches 793; Conservative 0; Mismatches 58; Indels 22; Gaps 15;

QY 287 GCGAGCCTGTGGCGGCTTCTGCGGCGCTACCGGAGAGGCGGCTGCGGCGGCGCTGC 346

```

Db 1 GGCAAGCCCTGTGCGCGCTTCCTCCGCGCTACCGCGAGGAGGAGCGCTGCGCGCGCTGCG 60
QY 347 AGAGAGCCTGTGCGCGCGCGCTTCGCGCGCTACCGCGAGGAGGAGCGCTGCGCGCGCTGCG 406
Db 61 AGAGAGCCTGTGCGCGCGCGCTTCGCGCGCTACCGCGAGGAGGAGCGCTGCGCGCGCTGCG 120
QY 407 CCGGCGCGCGAGCGCGC-TGAGCGCTTTGCTGCGCGAGCGAGAGCGCTGTTGAGTTGCATC 465
Db 121 CCGGCGCGCGAGCGCGCTGAGCGCTTTGCTGCGCGAGCGAGAGCGCTGTTGAGTTGCATC 180
QY 466 CTAGCCCGAGAGCGCGCGCGCTTCGCGCGAGCGAGAGCGCTGAGCGAGATGCGCTG 525
Db 181 CTAGCCCGAGAGCGCGCGCGCTTCGCGCGAGCGAGAGCGCTGAGCGAGATGCGCTG 240
QY 526 CGAAATCTGAAGTGCAGCGCTGCGGAGCGCGGAGTGCAGCGGAGAGTGCCTTCGCGCGCG 585
Db 241 CGAAATCTGAAGTGCAGCGCTGCGGAGCGCGGAGTGCAGCGGAGAGTGCCTTCGCGCGCG 300
QY 586 TTGCAAGCCCGCGGCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCGCGCGCGCGCG 645
Db 301 TTGCAAGCCCGCGGCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCGCGCGCGCGCG 359
QY 646 CAGACTTTTCTGTCCAGGAGTCAAGCTGTG-TGTGATCGCGCGCTGAGCGCTGAAGAGCA 704
Db 360 CAGACTTTTCTGTCCAGGAGTCAAGCTGTG-TGTGATCGCGCGCTGAGCGCTGAAGAGCA 419
QY 705 ACAGACGTTGCGCGCGCT-CTGTGAGTCTCAAAATGCGCGCAAGTGGAGCGCTCACTGAGC 763
Db 420 ACAGACGTTGCGCGCGCTCTGTGAGTCTCAAAATGCGCGCAAGTGGAGCGCTCACTGAGC 479
QY 764 GAGGCTGCGCGCGCGCTGCGGAGCGCGCGCGCTGAGCT--CGCTGCGCTACGATGAGCG 821
Db 480 GAGGCTGCGCGCGCGCTGCGGAGCGCGCGCGCTGAGCTTCCGCTGCGCTACGATGAGCG 539
QY 822 C-GAGGAGCTGTAGAGAGAGCGCTTCCAG-CTGCTGCGCGCGCTTCTGTGAGAGCGCGCG 879
Db 540 CGAGGAGAGCTGTAGAGAGAGCGCTTCCAGCTGTGCGCGCGCTTCTGTGAGAGCGCGCG 599
QY 880 CGCGCGCGCGAGCGCTGCGAGCGCGCTGTGAGAGCACTGAGG-AGAGAGAGCTCAACAGCT 938
Db 600 CGCGCGCGCGAGCTGAGCGCGCTGTGAGAGCACTGAGAGAGAGAGAGCTCAACAGCT 659
QY 939 -GGCAGAGAGCTTGTGCGCGCTGAGC--GATCCCAATGCGCGCGCTG-CCTAGAGCA-- 991
Db 660 GGGGAGAGAGCTGCTTGGGCTTGAACGATTCCCAATGGCGCTGAGCTGAGAGCAAG 719
QY 992 -GGGAGTGCAGCGAGCTTTTGAAGAGCTGATGCGCTTGAAGGTTCTTGTGCGGCTATTG 1050
Db 720 TGGGTTGCAAGCGAGTTTGGGAAACCTGATGCGCTTCAAGGTTCTTGTGCGGATATG 779
QY 1051 CTGAACCCCTGTCCATGCAAGGAGCGCTGAACCTGCACTTGAGCTATCTGTGAGAGCTGC 1110
Db 780 GCTGAACCTGTCCATGCAAGCGAGCTGAGAGCTGCACTTGAGCTATCTGTGAGAGCTGC 834
QY 1111 TGGGAGAGAGTTGATGCTTCCCGAGAGAGCA 1143
Db 835 CTGGGCAAGTGAATGCTTCCCGAGAGAGCA 867

Search completed: February 3, 2003, 16:12:57
Job time : 2225 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:24:48 ; Search time 66 Seconds
(without alignments)
6667.896 Million cell updates/sec

Title: US-09-763-748-1

Perfect score: 1435
Sequence: 1 cggcgccggtggaacca.....gataataagataacacg 1435

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1435	100.0	1435	US-09-143-212-1	Sequence 1, Appl1
2	1435	100.0	1441	US-08-414-625-1	Sequence 1, Appl1
3	427.6	29.8	1384	US-08-414-625-3	Sequence 3, Appl1
4	200	13.9	200	US-08-983-502-3	Sequence 3, Appl1
5	200	13.9	200	PCT-US96-10521-3	Sequence 3, Appl1
6	69.2	4.8	2793	US-08-209-747-1	Sequence 1, Appl1
7	69.2	4.8	2793	US-08-458-298-1	Sequence 1, Appl1
8	63.8	4.4	1926	US-09-249-585A-2	Sequence 2, Appl1
9	63.8	4.4	2580	US-09-050-863-2	Sequence 2, Appl1
10	63.8	4.4	2580	US-09-359-081-2	Sequence 2, Appl1
11	63.8	4.4	5452	US-09-130-114-1	Sequence 1, Appl1
12	63.8	4.4	9600	US-08-910-647-1	Sequence 1, Appl1
13	63.8	4.4	9600	US-09-620-925-1	Sequence 1, Appl1
14	63.8	4.4	10596	US-07-884-811-15	Sequence 15, Appl1
15	63.8	4.4	10596	US-07-885-971-15	Sequence 15, Appl1
16	63.8	4.4	10596	US-08-187-783A-15	Sequence 15, Appl1
17	63.8	4.4	10596	US-08-194-088A-15	Sequence 15, Appl1
18	63.8	4.4	10596	US-08-194-087-15	Sequence 15, Appl1
19	63.8	4.4	10596	PCT-US93-04648-15	Sequence 15, Appl1
20	55.6	3.9	9551	US-08-056-200-93	Sequence 93, Appl1
21	55.6	3.9	9551	US-08-800-644-93	Sequence 93, Appl1
22	53	3.7	1215	US-08-911-853-8	Sequence 8, Appl1
23	53	3.7	1215	US-09-479-409-8	Sequence 8, Appl1
24	53	3.7	1215	US-09-479-453-8	Sequence 8, Appl1
25	53	3.7	17612	US-08-911-853-29	Sequence 29, Appl1
26	53	3.7	17612	US-09-479-409-29	Sequence 29, Appl1
27	53	3.7	17612	US-09-479-453-29	Sequence 29, Appl1

28	51.4	3.6	1157	US-07-709-949-1	Sequence 1, Appl1
29	50	3.5	3241	US-09-434-288-11	Sequence 11, Appl1
30	48.4	3.4	4041	US-09-105-537-36	Sequence 36, Appl1
31	48.4	3.4	36778	US-09-105-537-5	Sequence 5, Appl1
32	48.2	3.4	4403765	US-09-103-840A-2	Sequence 2, Appl1
33	48.2	3.4	4411529	US-09-103-840A-1	Sequence 1, Appl1
34	48	3.3	516	US-09-183-861-34	Sequence 34, Appl1
35	48	3.3	516	US-09-022-765-34	Sequence 34, Appl1
36	48	3.3	3969	US-08-026-138E-16	Sequence 16, Appl1
37	47.8	3.3	43280	US-08-804-227C-1	Sequence 1, Appl1
38	47.4	3.3	43280	US-08-804-227C-1	Sequence 1, Appl1
39	46.8	3.3	38506	US-09-320-878-19	Sequence 19, Appl1
40	46.6	3.2	3489	US-08-728-323A-1	Sequence 1, Appl1
41	46.6	3.2	3489	US-09-298-568-1	Sequence 1, Appl1
42	46.6	3.2	32207	US-08-770-379-20	Sequence 20, Appl1
43	46.6	3.2	32207	US-08-757-659A-20	Sequence 20, Appl1
44	46.6	3.2	32207	US-09-230-371A-20	Sequence 20, Appl1
45	46.6	3.2	33529	US-09-144-085-3	Sequence 3, Appl1

ALIGNMENTS

```
RESULT 1
US-09-143-212-1
; Sequence 1, Application US//09143212B
; Patent No. 6077672
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia and Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRADD EXPRESSION
; FILE REFERENCE: RTS-0005
; CURRENT APPLICATION NUMBER: US/09/143,212B
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 87
; SEQ ID NO 1
; LENGTH: 1435
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(987)
; US-09-143-212-1

Query Match      100.0%; Score 1435; DB 3; Length 1435;
Best Local Similarity 100.0%; Pred. No. 5e+300;
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCGGCGGTGGGAACCCGAGCCCGCCGAGCGGCGAGAGTGAGATGGAGCTGGG 60
DB 1 CTGGCGGCGGTGGGAACCCGAGCCCGCCGAGCGGCGAGAGTGAGATGGAGCTGGG 60
QY 61 CAAATGGGCAACAAGATGGGTGGGACGCGATACCTGTTTGGAGTCTCTGGTGGAC 120
DB 61 CAAATGGGCAACAAGATGGGTGGGACGCGATACCTGTTTGGAGTCTCTGGTGGAC 120
QY 121 AAGGTGTCCTGTGCGATGCTTACGCGACCCCGACAGAGGTGACAGTTCACAGGGCT 180
DB 121 AAGGTGTCCTGTGCGATGCTTACGCGACCCCGACAGAGGTGACAGTTCACAGGGCT 180
QY 181 CTGACAGCTGCTTGGCAGAGAGCGGCGGAGCGCGAGCGTGTGAGATGCTGAAGATC 240
DB 181 CTGACAGCTGCTTGGCAGAGAGCGGCGGAGCGCGAGCGTGTGAGATGCTGAAGATC 240
QY 241 CACCGCAGGACCCCGAGCTGATCTGACGCTGCAATTCGCGGGGGGCAAGCCCTGGC 300
DB 241 CACCGCAGGACCCCGAGCTGATCTGACGCTGCAATTCGCGGGGGGCAAGCCCTGGC 300
QY 301 CGGTTCTCGGCGCTTACCGGAGGGGGGCGTGGCGCGCTGGAGAGAGGCGTGGG 360
DB 301 CGGTTCTCGGCGCTTACCGGAGGGGGGCGTGGCGCGCTGGAGAGAGGCGTGGG 360
QY 361 GCCGCGCTGCGCCAGCAGCTGCTGCTGCACTGAGACTGCGCGCGCGCGGCGGAGCGG 420
```

Db 361 GCCGCGCTCGCCAGCACTCGGTGCGCTGCAACTGAGCTGCGCGCGCCGAGCGG 420
Qy CTGAGCGCTTTGCTGCGGAGAGAGCGCTGTTGATGTTGATCTGATCCAGAGAGCC 480
Db CTGAGCGCTTTGCTGCGGAGAGAGCGCTGTTGATGTTGATCTGATCCAGAGAGCC 480
Qy 481 GACCGGCTCCGGGATGAGAACTGCTGAGCTGAGAGATCGCTGCGAATCTGAAGTGC 540
Db GACCGGCTCCGGGATGAGAACTGCTGAGCTGAGAGATCGCTGCGAATCTGAAGTGC 540
Qy 541 GCGTGGGCGCGCGGAGTGCAGCGGAGGTCGCTTGGCCCCCTTGCAGCGCCCGGTG 600
Db GCGTGGGCGCGCGGAGTGCAGCGGAGGTCGCTTGGCCCCCTTGCAGCGCCCGGTG 600
Qy 601 CCTCTGTCGAGAGTGAAGCGCGCGCGCGCCAGCTTGCCAGCTTTCTGTTT 660
Db CCTCTGTCGAGAGTGAAGCGCGCGCGCGCCAGCTTGCCAGCTTTCTGTTT 660
Qy 661 CAGGCTCAGCTGATGATGAGTGCAGCTGAGCTGAGAGCAACAGAGCTTTCGCGCC 720
Db CAGGCTCAGCTGATGATGAGTGCAGCTGAGCTGAGAGCAACAGAGCTTTCGCGCC 720
Qy 721 TCTGTGGCTTCAAAATGCGGCAAGGTGCGGCGCTCACTGAGCGAGCTGCGGCGCTG 780
Db TCTGTGGCTTCAAAATGCGGCAAGGTGCGGCGCTCACTGAGCGAGCTGCGGCGCTG 780
Qy 781 CGGAGCCCGGCGCTGAGACTCGCTGCGCTAAGATGAGCGCGGAGAGCTGTAACAGAG 840
Db CGGAGCCCGGCGCTGAGACTCGCTGCGCTAAGATGAGCGCGGAGAGCTGTAACAGAG 840
Qy 841 GCGTTCAGCTGCTGCGCGCTGCTGCGAGCGCGAGCGCGCGCGCGCGCGCTGAGCGC 900
Db GCGTTCAGCTGCTGCGCGCTGCTGCGAGCGCGAGCGCGCGCGCGCGCGCTGAGCGC 900
Qy 901 CTGTGTGAGGCACTGAGAGAGCAAGCTCACAGCTGCGAGAGAGCTTGTGCGCGCTG 960
Db CTGTGTGAGGCACTGAGAGAGCAAGCTCACAGCTGCGAGAGAGCTTGTGCGCGCTG 960
Qy 961 ACCGATCCCAATGAGGCGCTGCGCTGAGCAAGGAGTGCAGCGAGCTTTTGGAGAAC 1020
Db ACCGATCCCAATGAGGCGCTGCGCTGAGCAAGGAGTGCAGCGAGCTTTTGGAGAAC 1020
Qy 1021 ATGCGCTTAGGGTCTCTTCTGCGGCTAATGCTGAACCCCTGCTCAGCAAGAGCCCTGA 1080
Db ATGCGCTTAGGGTCTCTTCTGCGGCTAATGCTGAACCCCTGCTCAGCAAGAGCCCTGA 1080
Qy 1081 AACTCCACTTGGCTTATCTGCTGAGACTGCTGCGGCAAGTGTGATGCTTCCCGAGAG 1140
Db AACTCCACTTGGCTTATCTGCTGAGACTGCTGCGGCAAGTGTGATGCTTCCCGAGAG 1140
Qy 1141 CCAAGCACTGGGGGTGATCAATTGGGGATTCTGCTCAGGTAATTTGATTAAGTGTGG 1200
Db CCAAGCACTGGGGGTGATCAATTGGGGATTCTGCTCAGGTAATTTGATTAAGTGTGG 1200
Qy 1201 GTGGGGGGGAGCTTGTGTTGAGATCAGCTCAGCTTCTCCATCCCAAGAGGGGGCTTA 1260
Db GTGGGGGGGAGCTTGTGTTGAGATCAGCTCAGCTTCTCCATCCCAAGAGGGGGCTTA 1260
Qy 1261 CAGCAGCCCTTACAGTTTCACTCATGAAGCACTTGAATCTTGGTGTCTGAGCTTCA 1320
Db CAGCAGCCCTTACAGTTTCACTCATGAAGCACTTGAATCTTGGTGTCTGAGCTTCA 1320
Qy 1321 CCGTGGTGTCTGAGTATCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
Db CCGTGGTGTCTGAGTATCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
Qy 1381 TCACACTCAGCGTGAAGCCCGAATGTTAAAGCATGATTAATTAAGTGAAGCAG 1435
Db TCACACTCAGCGTGAAGCCCGAATGTTAAAGCATGATTAATTAAGTGAAGCAG 1435

RESULT 2
US-08-414-625-1

Sequence 1, Application US/08414625
Patent No. 5563039
GENERAL INFORMATION:
APPLICANT: Goedel, David V.
APPLICANT: Hsu, Hailing
TITLE OF INVENTION: INTRACELLULAR SIGNALING PROTEINS AND
METHODS OF USE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: FLEHR, HOENACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/414,625
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-60916/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1441 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 49..984
US-08-414-625-1
Query Match 100.0%; Score 1435; DB 1; Length 1441;
Best Local Similarity 100.0%; Pred. No. 5e-300;
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGGGCGGCGTGGGAAACCCAGGCCCCCGCGAGCGCGCCAGAGGTGAGATGGAGCTGGG 60
Db 1 CTGGGCGGCGTGGGAAACCCAGGCCCCCGCGAGCGCGCCAGAGGTGAGATGGAGCTGGG 60
Qy 61 CAAATGAGGACGAAGAGTGGGTGGGCAAGCGCATACCTGTTGTGAGATCCTGCTGGAC 120
Db 61 CAAATGAGGACGAAGAGTGGGTGGGCAAGCGCATACCTGTTGTGAGATCCTGCTGGAC 120
Qy 121 AAGTGTCTCTGTGAGATGCTTACGCGCAACCCCAAGAGAGTGGCAAGTGAAGAGGCT 180
Db 121 AAGTGTCTCTGTGAGATGCTTACGCGCAACCCCAAGAGAGTGGCAAGTGAAGAGGCT 180
Qy 181 CTGCAAGCTGCTTGGGAGAGAGCGGCGGAGCCCGAGCGTGTGAGATGCTGAAGATC 240
Db 181 CTGCAAGCTGCTTGGGAGAGAGCGGCGGAGCCCGAGCGTGTGAGATGCTGAAGATC 240
Qy 241 CACGCAAGCAGCCCGAGCTGATGTTGAGCTGAGCTTTCGCGGCGGAGCCCTGTGGC 300
Db 241 CACGCAAGCAGCCCGAGCTGATGTTGAGCTGAGCTTTCGCGGCGGAGCCCTGTGGC 300
Qy 301 CGCTTCTCTCGCGCTTACCGGAGAGGGGCGCTGCGCGCGGCTGCAAGAGAGCTTGGCG 360
Db 301 CGCTTCTCTCGCGCTTACCGGAGAGGGGCGCTGCGCGCGGCTGCAAGAGAGCTTGGCG 360
Qy 361 GCCGCGCTCGCCAGCACTCGGTGCGCTGCACTGAGAGCTGCGCGCGCGCGCGAGCGG 420

```

Db 361 GCCGCGCTCCGACGACTCGTGCCTGCACTGAGCTGCAGCGCCGAGCGG 420
Qy 421 CTGACGCTTTCGCGGGGCGAGAGAGCGCTTTTGAATTGCAATCCGACGACCC 480
Db 421 CTGACGCTTTCGCGGGGCGAGAGAGCGCTTTTGAATTGCAATCCGACGACCC 480
Qy 481 GACCGGCTCCGAGATGAAGAACTGCTGAGCTGAGAGATGGCTGCGAAATCTGAATGC 540
Db 481 GACCGGCTCCGAGATGAAGAACTGCTGAGCTGAGAGATGGCTGCGAAATCTGAATGC 540
Qy 541 GCGTCCGAGGCGCGGGGTCGAGCGGAGAGTCTTTCGCGCCCTTTCGAGCCCGGTC 600
Db 541 GCGTCCGAGGCGCGGGGTCGAGCGGAGAGTCTTTCGCGCCCTTTCGAGCCCGGTC 600
Qy 601 CCTCTCTGTCGAGAGTGAAGAGCGCCCGCCGCGCACTGCGCAGACTTTTCTGTC 660
Db 601 CCTCTCTGTCGAGAGTGAAGAGCGCCCGCCGCGCACTGCGCAGACTTTTCTGTC 660
Qy 661 CAGGGTCAGCTGTAGTGAATCGGCGCTGAGCTGAAGGACCAACAGACGTTGCGCGC 720
Db 661 CAGGGTCAGCTGTAGTGAATCGGCGCTGAGCTGAAGGACCAACAGACGTTGCGCGC 720
Qy 721 TCTGTGGGTCTCAAAATGCGCGCAAGTGGGCGCTCACTGCAAGAGCTGCGGCGCTG 780
Db 721 TCTGTGGGTCTCAAAATGCGCGCAAGTGGGCGCTCACTGCAAGAGCTGCGGCGCTG 780
Qy 781 CCGGACCCGCGCTGAGACTGCTGCGCTGAGAGTGAAGAGCGCGAGAGCTGTAGAGAC 840
Db 781 CCGGACCCGCGCTGAGACTGCTGCGCTGAGAGTGAAGAGCGCGAGAGCTGTAGAGAC 840
Qy 841 GCGCTTCAGCTGCTGCGGCGCTTGTGCGAGGCGAGGCGCGCGCGCGCTGCGAGCGC 900
Db 841 GCGCTTCAGCTGCTGCGGCGCTTGTGCGAGGCGAGGCGCGCGCGCGCTGCGAGCGC 900
Qy 901 CTGTGTGAGGCACTCGAGAGAGAGAGAGCTCAACGCTGCGAGAGAGCTTGTGTGAG 960
Db 901 CTGTGTGAGGCACTCGAGAGAGAGAGAGCTCAACGCTGCGAGAGAGCTTGTGTGAG 960
Qy 961 ACCGATCCCAATGCGGCGCTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 ACCGATCCCAATGCGGCGCTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Qy 1021 ATGCGCTTAAAGGCTTCTTTCGCGGCTATGCTGAACCCCTGTCATCCAGGAGCC 1080
Db 1021 ATGCGCTTAAAGGCTTCTTTCGCGGCTATGCTGAACCCCTGTCATCCAGGAGCC 1080
Qy 1081 AACTCCACTTGGCTTATCTGCTGAGAGCTGCTGCGGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 AACTCCACTTGGCTTATCTGCTGAGAGAGCTGCTGCGGAGAGAGAGAGAGAGAGAG 1140
Qy 1141 CCAGACCACTGCGGAGTGCATATGCGGAGATTCGCTCAGAGTCTTGTAGAGAGAG 1200
Db 1141 CCAGACCACTGCGGAGTGCATATGCGGAGATTCGCTCAGAGTCTTGTAGAGAGAG 1200
Qy 1201 GTGGGGGGGAGCTTGTGAGATCAGCTGACCTTCCATCCAGAGAGCGGGGCTTA 1260
Db 1201 GTGGGGGGGAGCTTGTGAGATCAGCTGACCTTCCATCCAGAGAGCGGGGCTTA 1260
Qy 1261 CAGCAGACCTTAACTGATTCATCATGAGAGACCTTGAATCTTGTGCTGAGAGAG 1320
Db 1261 CAGCAGACCTTAACTGATTCATCATGAGAGACCTTGAATCTTGTGCTGAGAGAG 1320
Qy 1321 CTTGGGTCGTCAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 CTTGGGTCGTCAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 TCACACTCAGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1435
Db 1381 TCACACTCAGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1435

```

```

US-08-414-625-3
; Sequence 3, Application US/08414625
; Patent No. 5563039
; GENERAL INFORMATION:
; APPLICANT: Goeddel, David V.
; APPLICANT: Hsu, Hailing
; TITLE OF INVENTION: INTRACELLULAR SIGNALING PROTEINS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLHR, HOEBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,625
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-60916/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1384 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..681
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..681
; US-08-414-625-3
Query Match 29.8%; Score 427.6; DB 1; Length 1384;
Best Local Similarity 78.5%; Pred. No. 1.9e-83;
Matches 540; Conservative 0; Mismatches 139; Indels 9; Gaps 2;
Qy 302 GCTTCTCCGCGCTACCGGAGGGGGGCTGCGCGCGCTGACAGAGAGCTGCGG 361
Db 2 GCTTCTCCAAAGCTTACCGGAGGGGGGCTGCGCACCGCGCTGACAGAGAGCTGCGG 61
Qy 362 CCGGCTCGCCAGACATCGGTGCTGCACTGAGAGCTGCGCGCGCGCGAGAGGCG 421
Db 62 CCGGCTTGGCCAGAGAGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121
Qy 422 TGACGCTTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
Db 122 TGACAGTTGGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 181
Qy 482 ACCGCTCCGAGATGAAGAACTGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 541
Db 182 ACCGCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 241
Qy 542 GCTCGGGGGCGCGGGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
Db 242 ACTGCACTGGCCAGGGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
Qy 602 CTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661

```

```

Db 302 CCTCTCCGACCGAGGAGAA-----ACCATCTGCGGCGCGCTGCGCAACTTTTCTGTTCC 355
QY 662 AGGGTCAGCCCTGTAGTGAATCGGCGCGCTGAGCCTGAAAGCAACAGACGTTTCGCGGCT 721
Db 356 ACGGCACTCGTAGTGAACCGGCCACTGACTCTTCAAGAACCAAGACAGCTGTTTCGCGCT 415
QY 722 CTGTGGGCTCTAAATAGGCGGCAAGGTGGGGGCGCTCACTGTCAGACGAGGCTGCCGGGCGCTGC 781
Db 416 CGGTGGGTCTCAAGTAGGCGGCGAGGGGTGGGGCGCTCGCTGCACTGTAATGTCGGGCACTGA 479
QY 782 GAGACCCGCGCGCTGGACTCTGCGCTGAGCTCAAGATACGAGCGGAGGAGACTGTAACGAGCAGG 841
Db 416 GAGATCTTGCTCTGCACTCGCTGCGCTTACGAGTATGAGCGCTGATGCGGCTATACAGACAGG 535
QY 842 CTTTCAGCTGCTGTCGAGCGCTTTCGTGACAGGCGGAGGCGCGCGGCAACGCTGACGCGCC 901
Db 536 CTTTCAGCTGCTGTCGCGCTTTCATGCAAGCGGAGGCGCGCGTGCACACTGACGCGCC 595
QY 902 TGTGTGAGGCACTCGAGAGAACGAGCTTCACAGCCTGCGC---AGAGACTTGTCTGGGCC 958
Db 596 TGTGTGAGGCGCTGAGAGAACGAACTCACTAGTCTAGCAGAGGAGATCTTGTGTGGCC 655
QY 959 TGACCGATCCCAATGCGGCGCGCTGAGCTA 986
Db 656 AGCGGAGCGGAGTGCGCGGCTGAGCTTA 683

RESULT 4
US-08-983-502-3
; Sequence 3, Application US/08983502
; Patent No. 639327
; GENERAL INFORMATION:
; APPLICANT: David WALLACH
; APPLICANT: Mark P. BOLDIN
; APPLICANT: Tanya M. GONCHAROV
; APPLICANT: Yuri V. GOLTSY
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
; TITLE OF INVENTION: AND OTHER PROTEINS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/983,502
; FILING DATE: 16-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10521
; FILING DATE: 14-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 114,615
; FILING DATE: 16-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 114,986
; FILING DATE: 17-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 115,319
; FILING DATE: 14-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 116,588
; FILING DATE: 27-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117,932
; FILING DATE: 16-APR-1996

```

ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25, 618
REFERENCE/DOCKET NUMBER: WALLACH-19
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
OS-08-983-502-3

Query Match	13.9%	Score 200;	DB 4;	Length 200;
Best Local Similarity	100.0%	Pred. No. 1.1e-34;		
Matches 200; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

[illegible]

```

RESULT 5
PCT-US96-10521-3
Sequence 3, Application PC/TUS9610521
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
TITLE OF INVENTION: AND OTHER PROTEINS
NUMBER OF SEQUENCES: 34
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10521
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 114,615
FILING DATE: 16-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 114,986
FILING DATE: 17-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 115,319
FILING DATE: 14-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 116,588
FILING DATE: 27-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117,932
FILING DATE: 16-APR-1996
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 base pairs
type: nucleic acid

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US96-10521-3

Query Match 13.9%; Score 200; DB 5; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 622 CGCCCGCCGCGCGCCACCTGCGCCAGACTTTTCTGTTCCAGGCTGACCTTGAGTAAT 681
DB 1 CGCCCGCCGCGCGCCACCTGCGCCAGACTTTTCTGTTCCAGGCTGACCTTGAGTAAT 60
QY 682 CGCCCGCTGAGCTGGAAGACCAACAGACGTTGCGCGCTCTGAGGCTCAATGCGC 741
DB 61 CGCCCGCTGAGCTGGAAGACCAACAGACGTTGCGCGCTCTGAGGCTCTCAATGCGC 120
QY 742 AAGTGGGCGCTCACTGACGAGGCTGCGGCGCTGCGGGAACCGGCGCTGACTCG 801
DB 121 AAGTGGGCGCTCACTGACGAGGCTGCGGCGCTGCGGGAACCGGCGCTGACTCG 180
QY 802 CTGGCCCTACGAGTACGAGCG 821
DB 181 CTGGCCCTACGAGTACGAGCG 200

RESULT 6

US-08-209-747-1
Sequence 1, Application US/08209747
Patent No. 5733771

GENERAL INFORMATION:

APPLICANT: Lewis, Randolph V.

TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider

TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider

NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolaach & Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-3487

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/209,747

FILING DATE: 14-MAR-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 1447-104P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-205-8000

TELEFAX: 703-205-8050

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2793 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ORGANISM: Nephila clavipes

TISSUE TYPE: minor ampullate gland

FEATURE:

NAME/KEY: CDS

LOCATION: 183..2675

OTHER INFORMATION: /product= "N. clavipes minor
OTHER INFORMATION: ampullate silk protein"
US-08-209-747-1

Query Match 4.8%; Score 69.2; DB 1; Length 2793;
Best Local Similarity 44.1%; Pred. No. 2.4e-06;
Matches 383; Conservative 0; Mismatches 478; Indels 7; Gaps 2;

QY 104 TGGAGTCTGCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 153
DB 1127 TGGAGTCTGCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1186
QY 164 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 223
DB 1187 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1246
QY 224 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 283
DB 1247 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1306
QY 284 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 343
DB 1307 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1362
QY 344 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 403
DB 1363 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1422
QY 404 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 463
DB 1423 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1482
QY 464 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 523
DB 1483 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1542
QY 524 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 583
DB 1543 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1602
QY 584 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 643
DB 1603 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1662
QY 644 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 703
DB 1663 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1722
QY 704 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 763
DB 1723 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1782
QY 764 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 823
DB 1783 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1842
QY 824 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 880
DB 1843 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1902
QY 881 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 940
DB 1903 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1962
QY 941 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 998
DB 1963 TGGAGTCTGGAACAAGGTGCTCTGTCGATGCTTACGCGACCCCGACAGAGG 1056

RESULT 7

US-08-209-747-1
Sequence 1, Application US/08458298
Patent No. 5756677

```

GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
TITLE OF INVENTION: CDNs Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESS: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,298
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,747
FILING DATE: 14-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
TISSUE TYPE: minor ampullate gland
FEATURE:
NAME/KEY: CDS
LOCATION: 183..2675
OTHER INFORMATION: /product= "N. clavipes minor
OTHER INFORMATION: ampullate silk protein"
US-08-458-298-1

Query Match 4.8%; Score 69.2; DB 1; Length 2793;
Best Local Similarity 44.1%; Pred. No. 2.4e-06;
Matches 383; Conservative 0; Mismatches 478; Indels 7; Gaps 2;
QY 104 TGGAGTCTGCTGTCGACAGAGTGTCTGCGATGCTTACGCGCAACCCGACAGAAAG 163
DB 1127 TGGAGTGTGCTGTTAGGTAAGAGTGTGCTGAGCTGAGCGCTGCAAGTGC 1186
QY 164 TGGCAGTGTACAGGGCTTCTGCAAGCTCTTGGCAGAGAGCGGCGGAGCGTGC 223
DB 1187 AGGACAGAGAGCTGAGAGGCTCAAGGTGTCAAAAGTATCGGTGCCGAGCAGAGAGCTGC 1246
QY 224 TGCAGATGTCTGAAGATCCACCGAGGACCGCGAGCTGATCGAGCTGCGATCTGCG 283
DB 1247 TGCAGTGTGAGAGAGAGAGCTGAGAGCGCTGTGTTACGTTAGAGTGTCTGTGCTGA 1306
QY 284 GCGCGAGCGCTGTGCGCCCTTCTTCGCGCTTACCCGAGAGGCGGCTGCGCGCGCGC 343
DB 1307 GAGAGAGCGCGCTGCGGTCTGTGAGAGAGAGCGCGTGC---GGGTGCAAGAGCTGAGAG 1362
QY 344 TGCAGAGAGAGCTGTGCGCGCGCGCTGCCAGACTCGGTGCCGTGCAACTGAGAGCTGC 403
DB 1363 CTACGATGTCAAGGTGGGTACGAGTGCAGTGTGAGAGAGCTGTGCGGTGCTGCTGCG 1422

```

```

QY 404 GCGCGCGCGCGAGCGCTGACCGCTTGTGCGGACAGAGAGCGCTGTTGAGTTGA 463
DB 1423 AGCAGAGAGCTGAGAGCGCTGTGTTACGGTAGAGGTGTGTCTGAGCTGAGCTGC 1482
QY 464 TCTTAGCCCAAGACCCGACCGCGCTCCGGGATGTAAGACTGTGCTGACTGAGAGATCCG 523
DB 1483 TGCAGGCGCAGAGAGCTGAGAGCTACGGTGTGCAAGGTGATACGGTCCGGAGAGAGC 1542
QY 524 TGCAGAACTGAAGTGCAGCTCGAGGAGCCCGGAGTGCAGAGGAGAGTCTTCGCGCC 583
DB 1543 TGTGCTGTCTACAGCTGTGTAACAGAGCCGAGCGCTGTGTGTTACGTAGAGGTGC 1602
QY 584 CTTGCAAGCCCGCGTGTGCTCTGTGCGAGTGAAGCCGCGCGCGCGCCACCTG 643
DB 1603 TGTGCTGTGAGCTGTGCTGCGCTGTGCGGAGGTGAGAGCAACGGTGTGTGATATGG 1662
QY 644 CCCAGACTTTTCTGTTCAGAGGTACGCTGTATGAATCGGCCCTGAGCTGAAGAGAC 703
DB 1663 TGCACAGCGCGTTATGTGTCCGAGAGAGAGCTGTGCGGCTGTCTGTGTGACAG 1722
QY 704 AACAGAGTTGCGCGCTCTGTGAGTCTCAATGCGCAAGTGGGCGCTCACTGCAGC 763
DB 1723 AGCAGAGAGTCTGTGTACCGTTAGAGTGTGTGTGAGCTGTGAGCTGTGAGAGTGC 1782
QY 764 GAGGCTGCGCGCGCTGTGCGGACCCGCGCTGAACTGCTGCGCTTACGATGAGCGCG 823
DB 1783 TGGAGCTGAGACCGCTGTGAGGTGACAGGAGAGCTGAGAGCTACGTTGTCAGGTGG 1842
QY 824 AGGAGCTGTACAGAGAGAGCTTCCAGCTGTGTGCGCGCTTGTGTCAAGGCC---GAGGCC 880
DB 1843 ATACGCTGCCGAGCAAGAGCTGTGTGCTGCGGAGCTGTGAGCAGAGAGCTGAGAGCGGC 1902
QY 881 GCCGCGCAGCGCTGAGAGCGCTGTGAGAGCACTGAGAGAGAGAGCTCACAGCTGAG 940
DB 1903 TCGGAGTTACAGTGAAGGTGTCTGTGAGAGAGCGCGTGTGTGTGAGTGTGAGCGCGCTGC 1962
QY 941 CAGAGAGCTTGTGCGCTGACCGATCC 968
DB 1963 AGGTCAAGAGAGAGAGCTGAGAGCTAC 1990

RESULT 8
US-09-249-585A-2
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent version 3.0
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

Query Match 4.4%; Score 63.8; DB 4; Length 1926;
Best Local Similarity 42.8%; Pred. No. 3.3e-05;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;
QY 103 GTGAGTCTCTGTGAGCAAGTGTCTGTGAGTGTCTGCGCAACCCGACAGAGAG 162
DB 260 GTGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319
QY 163 GTGCAAGTGTACAGGAGCTGTGAGAGCTGCTTGGCAAGAGAGCGGCGGAGAGCCGAGAGCTGTG 222

```

Db	320	GAGGGGCGAGGACAGAGAGAGGGGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCAG	379
Qy	223	CTGCAGATGTCTGAAGATCCACCGCAGCCGACCCTGATCTGTTCAGCTTGCAATTCTGC	282
Db	380	GAGGAGGGGCGAGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCAGAGAGAG	439
Qy	283	GGGCGGCGAGCCCTGTGGCCGCTTCTCTCCGCCCTTAACGCGGAGGGGGCGCTGCGCGCCG	349
Db	440	GGGCGAGGACAGAGAGAGAGGGGCGAGAGGGGCGAGAGCGCAGAGAGGGGCGAGAGGGGCGAG	499
Qy	343	CTGCAGAGAGACTTGCGCGGCGCGCTCCGCCAGCATCTGCTGCGCTGCAACTGAGACTG	402
Db	500	GAGGGGCGAGAGCAGAGAGAGGGGCGAGAGAGAGAGAGGGGCGAGAGGGGCGAGAGCAG	555
Qy	403	CGCGCGGCGCGAGCGGCGCTGAGCGCTTTGCTGCGGAGCGAGAGCGCTGTTGAGTTGC	462
Db	560	GAGGAGGGGCGAGAGAGGGGCGAGAGGGGCGAGAGCGCAGAGAGAGAGGGGCGAGAGAGGGG	619
Qy	463	ATTCCTAAGCCCGAGCGACCCGACCCGCTCCGGGATGAAGAATCTGCTGAGCTGAGAGATGCG	522
Db	620	CAGGAGGGGCGAGAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGAGGGGCGAGAGGGG	679
Qy	523	CTGCGAAATCTGAAGTGGGCTCGGGGGCCCCGGGGTGGCGAGCGGGAGAGTGTGCTTCGGCC	582
Db	680	CAGGAGGGGCGAGAGCAGAGAGGGGCGAGAGAGAGAGAGGGGCGAGAGAGGGGCGAGAG	739
Qy	583	CCCTTGACGCCCGCGTGCCCTCTCTGTGAGGTGAAGCCGCGCGCGCCGCGCACCT	642
Db	740	CAGGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGAGAGGGGCGAG	799
Qy	643	GCCCACTTTTCTGTTCTCAGGGCTACGCTTATGAATCGCGCGCTGAGCTTGAAGAGAC	702
Db	800	GAGGGGCGAGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGAGGGGCGAGAGAGGGG	858
Qy	703	CAACAGACGTTTCGCGCGCTCTGTGGGTCTCAAATGGCCCAAGTGGGGCGCTCACTGAC	762
Db	860	CAGGAGCGAGAGGGGCGAGAGGGGCGAGAGCGAGAGAGGGGCGAGAGGGGCGAGAGCAGAG	919
Qy	763	CGAGGCTGCGCGGGCGCTGCGGGAACCCGCGCTGGAATTCGCTGAGCTTAAGATAAGACGC	822
Db	920	GGGCGAGAGGGGCGAGAGCAGAGAGGAGCGCGAGAGCGAGAGGGGCGAGAGAGGAGGTG	979
Qy	823	GAGGACCTGTACGAGCAGGCGCTTCCAG	849
Db	980	GAGGCGGGGCTCGAGAGGCGAGTGGAG	1006
RESULT 9			
US-09-050-863-2			
Sequence 2, Application US/09050863			
Patent No. 611411			
GENERAL INFORMATION:			
APPLICANT: Lao, Ying			
APPLICANT: Hiang, Betty			
APPLICANT: Payan, Don			
TITLE OF INVENTION: Mammalian Protein Interaction Cloning			
NUMBER OF SEQUENCES: 5			
CORRESPONDENCE ADDRESS:			
ADDRESSSE: Flehr, Hohbach, Teet, Albritton & Herbert			
STREET: 4 Embarcadero Center, Suite 3400			
CITY: San Francisco			
STATE: CA			
COUNTRY: USA			
ZIP: 94111-4187			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patentin Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/050.863			

Query	DB	Score	Length	DB 3	Length	DB 3	Length
103	GTGAGTCTCGCTGAGCAAGTGGTCTCTGCGATGCTTACGCGCACCCAGCAGAG	162					
643	GTGAAACAGAGACAGAGACAGAGACAGAGACAGAGACAGAGACAGAGACAGAG	702					
163	GTGACAGTGTACAGAGCTCTGACAGCTGCTTGGCAGAGACCGCGGAGCCCGACGTG	222					
703	GAGGGGACAGAGACAG	762					
223	CTGCAGATGCTGAAGATCACCGACGCAACCCGACGCTGATCTGACCTGCAATTCTGC	282					
763	GAGGAGGGGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	822					
283	GAGGCGGACCCCTGCTGAGCGCTTCTCTCCGCGCTTACCGGAGAGGGGAGCTGCGCGCG	342					
823	GAGGACAGAGACAG	882					
343	CTGCAGAGAGACCTTGGCGCGCGCTGCGCTCCAGCATTGCTGCGCTGCAACTGAGAGCTG	402					
883	GAGGGGACAGAGACAG	942					
403	CGCGCGGCGCGGACAG	462					
943	GAGGAGGGGACAG	1002					
463	ATCTGACCCGACAG	522					
1003	CAGAGGGGACAG	1062					
523	CTGCGAATCTGAAGTGGGCTCGGGGGGCGGGGGGAGAGAGAGAGAGAGAGAGAGAGAG	582					
1063	CAGGAGGGGACAG	1122					
583	CCCTTGACAG	642					
1123	CAGGAGGGGACAG	1182					
643	GCCGCACTTTTCTGTTCCAGAGTCAAGCTTATGATTCGCGCGCTGAGCTGAGAGAGAC	702					
1183	GAGGGGACAG	1242					
703	CACAGACGTTTCGCGCGCTCTGTGGGTCTCAATATGCGCAAGTGGGGCGCTCACTGAG	762					
1243	CAGGAGGAG	1302					
763	CGAGGCTGCGGAG	822					
1303	GAGGAGGAG	1362					
823	GAGGAGCTGTACAG	882					
1363	GAGGAGGAG	1422					

RESULT 10
US-09-359-081-2
Sequence 2, Application US/09359081
Patent No. 6316223
GENERAL INFORMATION:
APPLICANT: Lao, Ying
Huang, Betty
Payan, Don
TITLE OF INVENTION: Mammalian Protein Interaction Cloning
System
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/359,081
FILING DATE: 22-Jul-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/050,863
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 949-8711
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2580 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2
Query Match 4.4%; Score 63.8; DB 4; Length 2580;
Best Local Similarity 42.8%; Pred. No. 3.5e-05;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 103 GTGAGTCTCGCTGAGCAAGGTGCTCTGTGATGCTTACGCGCAACCCCGACAGAA 162
DB 643 GTGGAACAGACAG 702
QY 163 GTGCAAGTGTACAGGCTCTGACAGGCTCTTGGCAGAGAGAGAGAGAGAGAGAG 222
DB 703 GAGGGGACAG 762
QY 223 CTGCAGATGCTGAAGATCCACCGACGACCCGACGATGATGATGATGATGATG 282
DB 763 GAGGAGGGGACAG 822
QY 283 GGGCGGACGCTGTGTGCGCTTCTCTCGGCTTACCGGAGAGAGAGAGAGAGAG 342
DB 823 GGGCAG 882
QY 343 CTGCAG 402
DB 883 GAGGGGACAG 942

QY 403 CGCCCGGCGCGAGCGGCTGTGACGCTTTGCTGCGGACGAGAGCCCTGTTGATTGC 462
DB 943 GAGAGAGGGGACAGAGAGGGGACAGAGAGGACAGAGAGAGAGAGAGAGAGAGAG 1002
QY 463 ATCTAGCCCGACAG 522
DB 1003 CAGAGAGGGGACAGAGAGGGGACAGAGAGGGGACAGAGAGAGAGAGAGAGAG 1062
QY 523 CTGCGAAATCTGAAGTGGCTCGGAGGAGCCGAGGAGGAGAGAGAGAGAGAGAG 582
DB 1063 CAGAGAGGGGACAGAGAGGGGACAGAGAGGGGACAGAGAGGGGACAGAGAGAG 1122
QY 583 CCTTGACGCCCCCGTGTCTCTGTGAGGTGAAGCCGCGCGCCGACCT 642
DB 1123 CAGAGAGGGGACAGAGAGGGGACAGAGAGGGGACAGAGAGGGGACAGAGAGAG 1182
QY 643 GCCAGACTTTTCTGTTCCAGAGGTACGCTGTATGATGAGGCGGCTGAGCTGAGAG 702
DB 1183 GAGGGGACAGAGAGAGAGAGGGGACAGAGAGGGGACAGAGAGGGGACAGAGAG 1242
QY 703 CAACAGAGCTTCGCGCGCTGTGTGAGTCTCAATGAGCGCAAGTGGGAGCTCACTG 762
DB 1243 CAGAGAGAGAGAGAGAGAGGGGACAGAGAGGGGACAGAGAGGGGACAGAGAGAG 1302
QY 763 CGAGGCTGCGGAGCGCTGTGAGGAGCCGCGCTGACTGCTGCTTACGATACGAGCC 822
DB 1303 GGGCAGAGAGGGGACAGAGAGAGAGAGAGGGGACAGAGAGGGGACAGAGAGAG 1362
QY 823 GAGGAGCTGTACGAGCAGGCTTCCAG 849
DB 1363 GAGGCGGGGTCGAGAGGCAAGTGGAG 1389

RESULT 11
US-09-130-114-1/c
Sequence 1, Application US/09130114
Patent No. 5976807
GENERAL INFORMATION:
APPLICANT: Horlick, Robert A.
APPLICANT: Damaj, Basam B.
APPLICANT: Robbins, Alan K.
TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
TITLE OF INVENTION: From Multiple Transfected Bpismoes
FILE REFERENCE: 0867/1D903US1
CURRENT APPLICATION NUMBER: US/09/130,114
CURRENT FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 5452
TYPE: DNA
ORGANISM: VEBNA
US-09-130-114-1
Query Match 4.4%; Score 63.8; DB 2; Length 5452;
Best Local Similarity 42.8%; Pred. No. 4e-05;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 103 GTGAGTCTCGCTGAGCAAGGTGCTCTGTGATGCTTACGCGCAACCCCGACAGAA 162
DB 2162 GTGGAACAGACAG 2103
QY 163 GTGCAAGTGTACAGGCTCTGACAGGCTCTTGGCAGAGAGAGAGAGAGAGAGAG 222
DB 2102 GAGGAGGGGACAG 2043
QY 223 CTGCAGATGCTGAAGATCCACCGACGACCCGACGATGATGATGATGATGATG 282
DB 2042 GAGGAGGGGACAG 1983
QY 283 GGGCGGACGCTGTGTGCGCTTCTCTCGGCTTACCGGAGAGAGAGAGAGAGAG 342
DB 1982 GGGCAG 1923

```

QY 343 CTGCAAGAGAGCTTGGCGCCGCGCTCGCCAGCACTGGTGGCTTGAATGAGCTG 402
Db 1922 GAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1663
QY 403 CGGCGCGCGCCGAGCGCGCTTGAAGCTTTTCTGGCGGAGAGAGAGAGAGAG 462
Db 1862 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1803
QY 463 ATCTGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 522
Db 1802 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1743
QY 523 CTGCAAAATCTGAAGTGGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582
Db 1742 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1683
QY 583 CCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 642
Db 1682 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1623
QY 643 GCCAGAGCTTTCTGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
Db 1622 GAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1563
QY 703 CAAAGAGCTTCCGCGCTCTGTGAGTCTCAAAATGCGCAAGTGGAGGCGCTCA 762
Db 1562 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1503
QY 763 CGAGGCTGCGGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 822
Db 1502 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1443
QY 823 GAGGAGCTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 849
Db 1442 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1416

```

RESULT 12

US-08-910-647-1

Sequence 1, Application US/08910647

Patent No. 6251433

GENERAL INFORMATION:

APPLICANT: Zuckermann et al.

TITLE OF INVENTION: Compositions and Methods for

TITLE OF INVENTION: Polynucleotide Delivery

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESS: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: California

COUNTRY: U.S.A.

ZIP: 94608-2916

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Releasee #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/910,647

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Fujita, Sharon M.

REGISTRATION NUMBER: 38,459

REFERENCE/DOCKET NUMBER: 1218.002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 923-2706

TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9600 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-910-647-1

```

```

Query Match 4.4%; Score 63.8; DB 4; Length 9600;
Best Local Similarity 42.8%; Pred. No. 4.4e-05;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

```

```

QY 103 GTGAGATCTCTGCTGAGCAAGTGTCTCTGTGAGTCTTACCGGACCCCGAGAG 162
Db 689 GTGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 748
QY 163 GTGAGAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 222
Db 749 GAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 808
QY 223 CTGCAAGTGTGAAGATCCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 282
Db 809 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 868
QY 283 GGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 342
Db 869 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 928
QY 343 CTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
Db 929 GAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 988
QY 403 CGGCGCGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462
Db 989 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1048
QY 463 ATCTTATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 522
Db 1049 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1108
QY 523 CTGCAAAATCTGAAGTGGCGCTTGGGGGCGCGGAGAGAGAGAGAGAGAGAGAGAG 582
Db 1109 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1168
QY 583 CCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 642
Db 1169 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1228
QY 643 GCCAGAGCTTTTCTGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
Db 1229 GAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1288
QY 703 CAAAGAGCTTCCGCGCTCTGTGAGTCTCAAAATGCGCAAGTGGAGCGCTTACAG 762
Db 1289 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1348
QY 763 CGAGGCTGCGGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 822
Db 1349 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408
QY 823 GAGGAGCTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 849
Db 1409 GAGGCGGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1435

```

RESULT 13

US-09-620-925-1

Sequence 1, Application US/09620925

Patent No. 6468986

GENERAL INFORMATION:

APPLICANT: Zuckermann et al.

TITLE OF INVENTION: Compositions and Methods for

TITLE OF INVENTION: Polynucleotide Delivery

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/620,925
FILING DATE: 21-Jul-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/910,647
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 1218.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9600 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1
Query Match 4.4%; Score 63.8; DB 4; Length 9600;
Best Local Similarity 42.8%; Pred. No. 4.4e-05;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;
QY 103 GTGAGTCTCTGCTGACAGAGGTGCTCTGTGGATGCTTACGCCACCCACAGAGAG 162
DB 669 GTGGAACAGAGACAG 748
QY 163 GTGACAGTGTACAGAGGCTCTGACAGGCTGCTTGAGAGAGAGAGAGAGAGAGAG 222
DB 749 GAGGGGACAGAGACAG 808
QY 223 CTGCAGATGCTGAAGATCCAGCCGACGACCCGACAGTGTATGTGACGCTGCAATTG 282
DB 809 GAGGAGGGGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 868
QY 283 GGGCGGCGAGCCCTGTGCGCGCTTCTCCCGCGGCTACCGGAGAGGGGGCGGTGCGCG 342
DB 869 GGGCAGAGACAG 928
QY 343 CTGCAGAGAGAGCTGAGCGCGCGCTCGCCAGCACTCGGTGCGCTGCACTGAGCTG 402
DB 929 GAGGGGACAGAGACAG 988
QY 403 CGCGCGCGCGAGCGGTGAGCGCTTTGCTGGCGGAGAGAGAGAGAGAGAGAGAG 462
DB 989 GAGGAGGGGACAG 1048
QY 463 ATCTTAAGCCACAGCCGACCGGCTCCGGAGTGAAGACTGGCTGAGCTGAGAGTGG 522
DB 1049 CAGAGAGGGGACAG 1108
QY 523 CTGCGAATCTGAAGTGCAGGCTCGAGGGGCGCGGGGTGCGAGCCGGAGAGGTGCTT 582
DB 1109 CAGGAGGGGACAG 1168
QY 583 CCTTTCAGCCCCCGGTGCTCTCTGTGTGAGAGTGAAGCCGCGCGCGCCCACT 642

DB 1169 CAGAGAGGGGACAGAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1228
QY 643 GCCCAGACTTTTCTGTTCCAGGGTCACTGTAAGTATCGCCGCTGAGCTTGAAGAC 702
DB 1229 GAGGGGACAGAGACAG 1288
QY 703 CAACAGAGCTTCCGCGCTCTGTGGTCTCAATGCGCAAGGTGGGGCCCTCAGCTCAG 762
DB 1289 CAG 1348
QY 763 CGAGGCTCCGCGGCGCTGCGGAGACCCGCGCTGAGACTGCTGGCTTACGATACGAGCG 822
DB 1349 GCGCAGAGAGGGGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408
QY 823 GAGGAGCTGTACAGAGAGCCCTTCAG 849
DB 1409 GAGGCCGGGCTCAGAGAGGCGAGTGGAG 1435
RESULT 14
US-07-884-811-15
Sequence 15, Application US/07884811
Patent No. 5316921
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mart, Melanie R.
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/884,811
FILING DATE: 19920518
CLASSIFICATION: 530
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 755.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10596 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-884-811-15
Query Match 4.4%; Score 63.8; DB 1; Length 10596;
Best Local Similarity 42.8%; Pred. No. 4.5e-05;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;
QY 103 GTGAGTCTCTGCTGACAGAGTGTCTCTGTGATGCTTACGGCGACCCACAGAGAG 162
DB 2223 GTGGAACAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2282
QY 163 GTGACAGTGTACAGGGCTCTGAGGCTGCTTGGCAGAGAGCGGCGAGAGCGTGG 222
DB 2283 GAGGGGACAGAGACAG 2342

QY 223 CTGCAGATGCTGAAGATCCACCGCAGCACCCTGATCTGTGACGTGCGATTCTGC 282
 Db 2343 GAGAGAGGGGCGAGGAGCAGAGGAGGGGCGAGAGGGGCGAGGAGGAGGAG 2402
 QY 283 GGGCGGAGCCCTGTGGCCCTTCTCTCCGGCTTACCGCAGAGGGGCGCTGGCCGGC 342
 Db 2403 GGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2462
 QY 343 CTGCAGAGGAGCCTGGGGGCGCGCTGGCCAGCCTGGGTCGCTGCACTGGAGCTG 402
 Db 2463 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2522
 QY 403 CGCGCCCGCCGAGCAGCGCTGAGCGCTTGTGCGGAGCAGGAGCAGCGCTTTGATTGC 462
 Db 2523 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2582
 QY 463 ATCTTACCCAGCAGCCCGACCGGCTCCGGGATGAAGACTGTGAGCTGAGGATGCG 522
 Db 2583 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2642
 QY 523 CTGCGAATCTGAAGTGGGCTGGGGGCGCGGGGTGGCCACCGGGGAGGTCTTGGCC 582
 Db 2643 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2702
 QY 583 CCTTTCAGGCCCCCGGTGCTCTGTGAGGAGTGAAGCCGCGCCGCGCCCACT 642
 Db 2703 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2762
 QY 643 GCCCAGACTTTTCTGTTCCAGGAGTCAAGCTGTGATGAGCCGCTGAGCCTGAAGGAC 702
 Db 2763 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2822
 QY 703 CAACAGAGCTTTCGCGGCTCTGTGGGTCTCAATGCGCAAGGGGGCGCTCACTGAC 762
 Db 2823 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2882
 QY 763 CGAGGCTGCGGGGCGCTGCGGAGCAGCGGCTGAGACTGCTGAGCTTGAAGGAC 822
 Db 2883 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2942
 QY 823 GAGGAGCTGTACGAGCAGCGCTTCCAG 849
 Db 2943 GAGGCGGGGTGAGGAGGAGGAGGAG 2969

RESULT 15
 US-07-885-971-15
 / Sequence 15, Application US/07885971
 / Patent No. 5328837

/ GENERAL INFORMATION:
 / APPLICANT: Godowet, Paul J. Lokker, Nathalie A. Mark, Melanie R.

/ TITLE OF INVENTION: HEPATOCTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
 / NUMBER OF SEQUENCES: 21
 / CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Genentech, Inc.
 / STREET: 460 Point San Bruno Blvd
 / CITY: South San Francisco

/ STATE: California
 / COUNTRY: USA

/ ZIP: 94080
 / COMPUTER READABLE FORM:

/ MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 / COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: patin (Genentech)

/ CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/07/885,971

/ FILING DATE: 19920518
 / CLASSIFICATION: 530

/ PRIOR APPLICATION DATA:
 / APPLICATION NUMBER:

/ FILING DATE:

/ ATTORNEY/AGENT INFORMATION:
 / NAME: Dieger, Ginger R.
 / REGISTRATION NUMBER: 33,055
 / REFERENCE/DOCKET NUMBER: 779
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 415/225-3216
 / TELEFAX: 415/952-9881
 / TELEX: 910/371-7168
 / INFORMATION FOR SEQ. ID NO: 15:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 10596 bases
 / TYPE: NUCLEIC ACID
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / US-07-885-971-15

Query Match 4.4%; Score 63.8; DB 1; Length 10596;
 Best Local Similarity 42.8%; Pred. No. 4.5e-05;
 Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 103 GTGAGATCTCTGCTGAGCAAGTGTCTGTCTGGATGCTTACGCGCAGCACCCTGAGGAG 162
 Db 2223 GTGAGACAGGAGCAGAGCAGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 2282
 QY 163 GTGAGATGTACAGGGCTGTGACAGCTGCTTGGCAGAGAGCGGCGGAGCCTCGACGTG 222
 Db 2283 GAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2342
 QY 223 CTGCAGATGCTGAAGATCCACCGCAGGAGCCGAGCTGATCTGCACTGCGATTCTGC 282
 Db 2343 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2402
 QY 283 GGGCGGAGGCTGTGGGCGCTTCTCCGCGCTTACCGGAGGGGCGCTGGCGCGCG 342
 Db 2403 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2462
 QY 343 CTGCAGAGGAGCCTTGGCGGCGCTGCGCCAGCAGCTCGGTGCGCTGCACTGAGGAGTGC 402
 Db 2463 GAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2522
 QY 403 CGGCGCGGCGCCGAGCGGCTGTGACGCTTGTGCGGAGTGAAGCGCGCGCGCCCACT 462
 Db 2523 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2582
 QY 463 ATCTTACCCAGCAGCCCGACCGGCTCCGGGATGAAGACTGTGAGCTGAGGATGCG 522
 Db 2583 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2642
 QY 523 CTGCGAATCTGAAGTGGGCTTGTGGGTCTCAATGCGCAAGGGGGCGCTCACTGAC 582
 Db 2643 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2702
 QY 583 CCTTTCAGGCCCCCGGTGCTCTGTGAGGAGTGAAGCGCGCGCGCGCCCACT 642
 Db 2703 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2762
 QY 643 GCCCAGACTTTTCTGTTCCAGGAGTCAAGCTGTGATGAGCCGCTGAGGAGGAC 702
 Db 2763 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2822
 QY 703 CAACAGAGCTTTCGCGGCTCTGTGGGTCTCAATGCGCAAGGGGGCGCTCACTGAC 762
 Db 2823 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2882
 QY 763 CGAGGCTGCGGGGCGCTGCGGAGCAGCGGCTGAGACTGCTGAGCTTGAAGGAC 822
 Db 2883 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2942
 QY 823 GAGGAGCTGTACGAGCAGCGCTTCCAG 849
 Db 2943 GAGGCGGGGTGAGGAGGAGGAGGAG 2969

Tue Feb 4 11:19:01 2003

us-09-763-748-1.rni

Page 12

Search completed: February 3, 2003, 16:14:31
Job time : 89 secs

...

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:25:18 / Search time 80 Seconds

(without alignments)
8058.688 Million cell updates/sec

Title: US-09-763-748-1

Sequence: 1 cggcgcgggcggtggaccacca.....gataataagatataacacg 1435

Scoring table: IDENTITY_NIC

Gapop 10.0, Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Published Applications NA:
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
3: /cgn2_6/prodata/2/pubpna/US05_PUBCOMB.seq:*
4: /cgn2_6/prodata/2/pubpna/US04_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US03_PUBCOMB.seq:*
6: /cgn2_6/prodata/2/pubpna/US02_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US01_PUBCOMB.seq:*
8: /cgn2_6/prodata/2/pubpna/US00_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145.4	10.1	325	10	US-09-960-352-12660
2	56.2	3.9	2108	10	US-09-962-832-225
3	55.4	3.9	14800	10	US-09-954-456-1601
4	52	3.6	1156	10	US-09-827-854-11
5	51.4	3.6	1157	10	US-09-954-456-760
6	51.4	3.6	1157	10	US-09-880-107-2244
7	51.2	3.6	41907	10	US-09-967-013-5
8	50.4	3.5	1156	9	US-09-870-759-129
9	50.4	3.5	1156	10	US-09-827-854-7
10	50.4	3.5	1156	10	US-09-827-854-8
11	50.4	3.5	1156	10	US-09-827-854-10
12	50.4	3.5	1156	10	US-09-827-854-12
13	50.4	3.5	1291	12	US-10-044-090-454
14	50.4	3.5	4826	10	US-09-774-304A-1
15	50.2	3.5	88421	9	US-09-976-059-1
16	49.6	3.5	1614	9	US-09-976-740-45
17	49.6	3.5	1614	12	US-10-023-529-45
18	49.6	3.5	1614	12	US-10-023-523-45
19	49.6	3.5	12425	9	US-09-976-740-50

C	20	49.6	3.5	12425	12	US-10-023-529-50	Sequence 50, Appl
C	21	49.6	3.5	12425	12	US-10-023-523-50	Sequence 50, Appl
C	22	49	3.4	2176	9	US-10-032-159A-19	Sequence 19, Appl
C	23	48.8	3.4	1156	10	US-09-827-854-9	Sequence 9, Appl
C	24	48.8	3.4	2561	9	US-09-976-740-48	Sequence 48, Appl
C	25	48.8	3.4	2561	12	US-10-023-529-48	Sequence 48, Appl
C	26	48.8	3.4	2561	12	US-10-023-523-48	Sequence 48, Appl
C	27	48.4	3.4	4041	9	US-09-860-846-36	Sequence 36, Appl
C	28	48.4	3.4	4041	10	US-09-861-289-36	Sequence 36, Appl
C	29	48.4	3.4	36778	9	US-09-860-846-5	Sequence 5, Appl
C	30	48.4	3.4	36778	10	US-09-861-289-5	Sequence 5, Appl
C	31	48	3.3	516	9	US-09-991-496-34	Sequence 34, Appl
C	32	48	3.3	516	10	US-09-874-923-34	Sequence 34, Appl
C	33	48	3.3	582	9	US-09-991-496-114	Sequence 114, Appl
C	34	48	3.3	582	10	US-09-874-923-114	Sequence 114, Appl
C	35	48	3.3	7065	9	US-09-991-496-115	Sequence 115, Appl
C	36	48	3.3	7065	10	US-09-874-923-115	Sequence 115, Appl
C	37	47.8	3.3	2561	12	US-10-023-529-48	Sequence 48, Appl
C	38	47.8	3.3	2561	12	US-10-023-523-48	Sequence 48, Appl
C	39	47.8	3.3	2561	12	US-10-023-523-48	Sequence 48, Appl
C	40	47.2	3.3	1233	9	US-09-712-363-77	Sequence 77, Appl
C	41	47	3.3	3345	10	US-09-894-998-49	Sequence 49, Appl
C	42	46.6	3.2	4366	12	US-10-044-090-668	Sequence 668, Appl
C	43	45.6	3.2	11220	9	US-09-860-846-32	Sequence 32, Appl
C	44	45.6	3.2	11220	10	US-09-861-289-32	Sequence 32, Appl
C	45	45.4	3.2	457	10	US-09-960-352-12765	Sequence 12765, A

ALIGNMENTS

RESULT 1
US-09-960-352-12660
; Sequence 12660, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Bryant, John C.
; APPLICANT: Mathiasagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12660
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 54-LIB3058-020-Q1-K1-F6
US-09-960-352-12660

Query Match 10.1%; Score 145.4; DB 10; Length 325;
Best local similarity 86.4%; Pred. No. 2, 7e-24;
Matches 172; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
QY 1 CTGCGGCGCTGGGACCCAGGCCCCCGAGGCGGCGGAGGTGATGCGAGCTGG 60
DB 22 CCGCGGCGCTGGGACCCAGGCCCCCGAGGCGGCGGAGGTGATGCGAGCTGG 80
QY 61 CAAATGCGACGAGAGGTGGGCGGAGGCGGATCTGTTTGGAGCTCTCGTGGAC 120
DB 81 CCAATGGGCTTGAAGGTGGGCGGAGGCGGATCTGTTTGGAGCTCTCACTGAC 140
QY 121 AAGGTGCTCTGTCGATGCTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCT 180
DB 141 AAGGTGCTCTGTCGATGCTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCT 200
QY 181 CTGCGGCGCTGCTGGGAG 199
DB 201 CTCGAGACTGCCCTAGCAG 219

```
RESULT 2
US-09-962-832-225
; Sequence 225, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 225
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-225

Query Match          3.9%; Score 56.2; DB 10; Length 2108;
Best Local Similarity 45.0%; Pred. No. 0.0007;
Matches 303; Conservative 0; Mismatches 358; Indels 12; Gaps 2;

QY 255 GCACTGATCTGTCAGCTGCGATTCTTCCGCGCGAGCCCTGTGGCCGCTTCTCCGCGC 314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 620 GCAGCAGGAGGGGCGAGCTGAGCTCCACAGCAGCAGGAGGGGCGAGCTGAGCTCCACAG 679
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 315 CTAACCGCGAGGGGGGCTGCGCGCGCTGCAAGAGGAGCCCTGGGCGCGCCCTGCGCCA 374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 660 GCAGCAGGAGGGGCGAGCTGAGCTCCACAGCAGCAGGAGGGGCGAGCTGAGCTCTGTA 739
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 375 GCACTGAGTCCGCTGCACTGAGAGCTGCGCGCGCGCGCGAGCGGCTGAGAGCTTGTCT 434
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 740 GCA----GCAGAGGGGCGAGCTGAGCTCTCTGAGCAGCAGAGGGGCGAGCTGAGAGTCTC 796
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 435 GCGCGACGAGAGCCGCTGTTTGAATTGCATCTTACGCCACAGCCCGACCGGCTCCGGGA 494
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 797 TGAGCAGCAGGAGGAGCAGCTGAAAGCAGCTGAGCAGCAGGAGGGGCGAGCTGAGAGTCCC 856
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 495 TGAAGACAGCTGAGCTGAGAGATGCGCTGGAATCTGAAGTGTGCGGCGCGCGCG 554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 857 AGAGGAGCAGATGGGCGAGCTGAAGTACTTGAAACAGCAGGAGGGGCGAGCTGAAGCAGCT 916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 555 GGGTGGCGACGGGCGAGGTGCTTCCGCGCGCTTGGCAGCCCGGTTGCCCTCTCTGTGGA 614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 917 GGATCAGCAGGAGGAGGAGCCAGAGCTCCACAGCAGCAGATGGGGCGAGCTGAAGCAGCT 976
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 615 GGTGAAGCCGCGCGCGCGCGCCACCTGCGCAGACTTTCTGTTCAGAGGTGAGCTGCT 674
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 977 GGAGCAGCAGGAGGGGCGAGCTTAAAGATCTGAGCAGCAGGAGGGGCGAAGAGAGAGCT 1036
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 675 AGTGAATGGCGCGCTGAGCTGAAGGACCAAGACGTTTGGCGCGCTGTGTGGGTCTCA 734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1037 GAGGAGCAGGAGGGGCGAGCTGAAGCAGCTGAGCAGCAGGAGGGGCGAGCTGAAGAGCT 1096
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 735 ATGCGCAGAGTGGGGCGCTCACTGACAGCAGAGCTGCCGGGCGCTGCGGAGCCCGGCGCT 794
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1097 GAGAGCAGCAGGAGG-----GCAAGCTGGGGGCTCCAGAGCAGCAGAGTGTGTAAGCT 1147
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 795 GGAATCGCTGCGCTTACGAGTACGAGCGCAGGAGGAGCTGTACGAGCAGGCGCTTCCAGTGTCT 854
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1148 GAAGCAGCTAGAGAGCAGCAGGAGGAGCAGCAAGCAGCCTGAGAGAGAGAGAGAGGAGAGCT 1207
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 855 GCGGCGCTTCTGTCAGGAGCGGAGAGGCGCGCGCGCAGCGCTGACGCGCTGTGTGAAGCAGCT 914
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1208 GAAGCAGCTGTGTCAGCAGGAGGGGCGAGCTGAAGCATCTGTGTGTCAGCAGGAGGAGGAGCT 1267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 915 CGAGAGAGAGAG 927
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
DB 1268 GAGCAGCAGAGAG 1280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-09-954-456-1601
; Sequence 1601, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cance
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1601
; LENGTH: 14800
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1601

Query Match          3.9%; Score 55.4; DB 10; Length 14800;
Best Local Similarity 43.0%; Pred. No. 0.0017;
Matches 382; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 30 GAGGCGGCGAGAGGTGAGATGCGAGCTGGGCAAAATGGGACGAAAGAGTGGTGGGAG 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4662 GCGGCGAGGCGAGAGGTGAGAGCTGGGCTGCGGCTGTAAGGCGGAGACCGAGCGGCGCGGA 4721
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 90 CGCATACTGTTTGTGAGAGTCTGCTGTGACAGAGTGTCTGTTCGAGATGCTTACGCGCA 149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4722 GAAGCAGAGGGGCGCTGCGAGGCGCTGGAAGAGCTGCGGCTGAGGCGGAGGCGGACG 4781
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 150 CCCCAGAGAAAGTGGAGAGTACAGGGCTCTGAGAGGCTGCGCTTGGC---AGAGAGCGG 206
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4782 GCGCTTGGCGAGCGCGAGGTGAGACGAGCGCGGAGGTACAGGTGAGGCTTGTGAGAGCGG 4841
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 207 CGGAGCCCGGACGTGCTGAGATGCTGAAGATCCACCGCAGCAGCCGCGAGCTGATGCT 266
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4842 GCAGGCGAGTGAAGAGGCGGAGGCTGCAAGACAAAGCGCTCTTTCGCGAGAGAGAGCGG 4901
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 267 GCAGCTGCAATTTCTGCGGGGCGAGCCCTTGTGGCCGCTTCTCCGCGGCTTCCCGAGAGG 326
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4902 ACAGCTGAGAGCGCTCCCTGCGAGGAGGAAACAGTGTGTGTGCAACAGCTGCGGAGAGAGG 4961
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 327 GCGCGTGGCGCGCGCGCTGCGAGAGGAGCCTGCGCGCGCGCTGCGCGCAGCAGCTGAGTGC 386
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4962 TGAAGCGGCGGAGCAGCAGGAGCGGAGGCGGAGGCGGCGGCGAGAGAGAGAGAGGAG 5021
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 387 GCTGCACTGAGAGCTGCGCGCGCGCGCGAGCGGCTGAGAGCTTGTGCTGCGGAGAGAG 446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
Db 5022 GCTGAGCGCTGAGCTGAGGCAAGGCGCTACGCTGCGCTGAGCGGAGGA 5081
Qy 447 GCGCTGTTTGAAGTTCATCTTACGCGCCAGCAGCCCGGCTCCGGAGTGAAGACTGAC 506
Db 5082 GGTGGCCGAGAGAGAGAGCTTGGC---GCAAGCCGAGAGCTTAAGAGCAAGAGAGAGG 5138
Qy 507 TGAGCTGAGAGATGCTGCTGCGAAATCTGAAGTGCAGCTCGGAGGAGCCCGGAGTGCAG 566
Db 5139 GGAAGCGAGAGCGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 5198
Qy 567 GGAAGTGCCTTGGAGCCCTTGGAGAGCCCGGCTCTTCTGTGAGAGTGAAGCCGCC 626
Db 5199 GGCCTGAACAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 5258
Qy 627 GCGCGCGCGCGCACCTGCGCAGACTTTTCTGTCCAGAGGCTGAGGCTGAAGTGAATTCG 686
Db 5259 GCGCGCGAGAGAGAGAGTGAATTCGAGCTGCGGCGCGAGAGAGAGAGAGAGAGAGAG 5318
Qy 687 GCTGAGCTTGAAGAGCAAGAGAGTTCGCGCGCTCTGTGAGTCTCAATGAGCGAAGT 746
Db 5319 GCAGCTGCTGAGAGAGAGAGTGGCCCGGCTGAGAGTGAAGAGAGAGAGAGAGAGAA 5378
Qy 747 GGGCGCTCATCTGAGAGAGAGCTGCGGAGCGCTGCGGAGAGAGAGAGAGAGAGAGT 806
Db 5379 ACGGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 5438
Qy 807 CTAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 866
Db 5439 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 5498
Qy 867 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 915
Db 5499 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 5547
```

```
RESULT 4
US-09-827-854-11
; Sequence 11, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zambis, Vasiliis
; APPLICANT: Kyriacs, Kyriacos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-11
```

```
Query Match 3.6%; Score 52; DB 10; Length 1156;
Best Local Similarity 44.2%; Pred. No. 0.0053;
Matches 214; Conservative 0; Mismatches 270; Indels 0; Gaps 0;
```

```
Qy 108 GTCTCTGCTGAGCAAGAGTGTCTGTGAGATGCTTACGCGCAGCCCGCAGAGAGTGGC 167
Db 360 GACCCCGGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 419
Qy 168 AGTGTACAGAGGCTTGCAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 227
Db 420 CCGGCTGGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 479
Qy 228 GATGCTGAAGATCCAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 287
```

```
Db 480 GCAAGCCATCTCGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 539
Qy 288 GCAAGCCATCTCGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 347
Db 540 CAGCTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 599
Qy 348 GAGAGAGCTGCGGCGCGCTGCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 407
Db 600 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 659
Qy 408 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 467
Db 660 GCGCTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 719
Qy 468 AGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 527
Db 720 GCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 779
Qy 528 AAGTCTGAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 587
Db 780 CAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 839
Qy 588 GCAG 591
Db 840 GCTG 843
```

```
RESULT 5
US-09-954-456-760
; Sequence 760, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 760
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-760
```

```
Query Match 3.6%; Score 51.4; DB 10; Length 1157;
Best Local Similarity 46.5%; Pred. No. 0.0073;
Matches 166; Conservative 0; Mismatches 191; Indels 0; Gaps 0;
```

```
Qy 235 AAGATCCAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 294
Db 488 ATGCTCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 547
```

QY	295	TGTGGCGGCTTCTCTCCGGCGCTTACCGGAGGGGGGGCGTGGCGCGCGGCGTGCAGAGAGC	354
Db	548	CGTAAAGCGGCTCTTCCGGATCCGGATACCTGCAGAAAGCCCTGGAGAGTGTACAGCC	607
QY	355	CTGGCGGCGCGGCTTCCGCCAGCACTCGGTCCGGCTGCMACTGAGAGTCGCGCGCGGCGCC	414
Db	608	GGGGGCGCGGCGCGGCGCGGCGCTCAAGCCCACTCCGCGAGCGCCTTGGAGGCCCTTG	667
QY	415	GAGCGGCGTGCAGCGCTTTCCTGGCGGAGAGGAGGCTGTTTGATTTGCATCTTACGCCAG	474
Db	668	GTCGAACAGGGGCGCGGTGCGGAGCCGCCACTGTGGCTTCTTGGCGGCGCAGCGGCTACAG	727
QY	475	CAGCCCGACCGGCTTCCGGATGAAGAACTGGCTAGCTTGAAGATGCGCTGCAGAAATCTG	534
Db	728	GAGCGGCGCTCAAGGCTCTGGGGGCGAGCGGCTTCGCGCGCGGAGTGAGAGATGGGACGCGG	787
QY	535	AAGTCCGGCTCGGGGGCGCGGGGTGGCGACCGGGAGATCGCTTGGGCGCCCTTTCGAG	591
Db	788	AACCGCGACCGCTTGCAGAGATGAAGAGCAGGTGGCGGAGGTCGCGCGCAAGCTG	844

RESULT 6
 US-09-880-107-2244
 Sequence 2244, Application US/09880107
 Patent No. US20020142981A1
 GENERAL INFORMATION:
 APPLICANT: Horne, DarcI T.
 APPLICANT: Vockley, Joseph G.
 APPLICANT: Scherf, Uwe
 APPLICANT: Gene Logic, Inc.
 TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 FILE REFERENCE: 44521-5028-WO
 CURRENT APPLICATION NUMBER: US/09/880,107
 CURRENT FILING DATE: 2001-06-14
 PRIOR APPLICATION NUMBER: US 60/211,379
 PRIOR FILING DATE: 2000-06-14
 PRIOR APPLICATION NUMBER: US 60/237,054
 PRIOR FILING DATE: 2000-10-02
 NUMBER OF SEQ ID NOS: 3950
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2244
 LENGTH: 1157
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Genbank Accession No. US20020142981A1 MI2529
 US-09-880-107-2244

Query Match	3.6%	Score 51.4;	DB 10;	Length 1157;
Best Local Similarity	46.5%	Pred. No. 0.0073;		
Matches 166;	Conservative 0;	Mismatches 191;	Indels 0;	Gaps 0

Qy 535 AAGTCCGCTTGGGGGCCCCGGGGAGTCCGTTCCGCCCCCTTGCAG 591
Db 788 ACCCGCGACCGCTTGCAGAGTGAAGAGCAGCTGGCGGAGTGGCGCCAACTG 844

```

RESULT 7
US-09-967-013-5
: Sequence 5, Application US/09967013
: Patent No. US20020045840A1
: GENERAL INFORMATION:
: APPLICANT: Staton, Jr. Vincent P.
: TITLE OF INVENTION: METHOD FOR GENETIC AN
: TITLE OF INVENTION: AND USES THEREOF
: FILE REFERENCE: 11926-022001
: CURRENT APPLICATION NUMBER: US/09/967,013
: PRIORITY FILING DATE: 2000-10-25
: PRIOR APPLICATION NUMBER: 60/206,613
: PRIOR FILING DATE: 2000-05-23
: NUMBER OF SEQ ID NOS: 91
: SOFTWARE: FASTSQ for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 41907
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-967-013-5

```

Query Match	3.6%;	Score 51.2;	DB 10;	Length 41907;
Best Local Similarity	43.8%;	Pred. No. 0.02;		
Matches 212;	Conservative 2;	Mismatches 270;	Indels 0;	Gaps 0;

QY	108	GTTGCGCTGGAGCAAAAGTGTCTCTGTGGATGCTTACGCGACACCCGACGAAAGGTGGC	167
Db	21162	GACCCCGGTGTGGCGAAGAACCGCGGACCGGCTGTCCAAAGAGTTGAGGGCGCGCAGAGC	21221
QY	168	AGTGTACAGAGGCTCTGTACAGGCTGCCTTTGAGCAGAGGCGCGGAGCCCGSACGTGTCTGA	227
Db	21222	CCGGCTGTGGCGGAGACATGAGAGACGTGTGGCGCCGCTGTGTGACATGACCGCGGACAGT	21281
QY	228	GATGTGTAAGATTCACCCGACAGCGACCCGACGTGATGTGTGACGTGTGATTTCTGCGGAGC	287
Db	21282	GCAAGGCATCTCTGGCCTAAGACACCGAAGACCTGGGGTGGGCTCGCTCCACCTGGG	21311
QY	288	GCAACCTGTGGCCGCTTCTCTCCGCGCTTACCGGAGAGGGGCGCTGTGGCGCCGCGCTGCA	347
Db	21342	CAAGCTGTGTAAGCGGCTCTCTCCGAGTCCGATGACCTGTCAAGAGTGCCTGTGCACTGTA	21401
QY	348	GAGAGCTGTGGCGCGCGCGCTGTGCCACACACTCGGTCCGCTGCACACTGTGAGCTGTGGCGC	407
Db	21402	CCAGCGCGGGGCGCGGAGAGGGCGCGACCGGCGCTTACGCGCCATCCGCGCAGCGCTTGG	21461
QY	408	CGGCGCGAGCGGCTTGTGACGCTTTTCTGTGGCGAGACGAGAGACGCTGTGTTGATGTGCATCT	467
Db	21462	GCCCCGTGTGTGAACAAGGGCGCGGTGTGGGGCCGCACTGTGTGGCTCCTGTGGCGGCGACGC	21521
QY	468	AGCCACAGACCCGACCGGCTTCGGGATGAAAGAACTGTGACTGTGAGAGTACGCTTGGC	527
Db	21522	GCTTACAGAGCGGGCCCAAGGCTGTGGGGGAGACGGGCTGTGGCGCGGAGTGAAGAGATGGG	21581
QY	528	AAATCTGAAGTGGGCTCGGGGGGCCCGGGGTGGCGACGGGAGGTGTGCTTTCGCGCCCTTT	587
Db	21582	CAGCGGACCCGCGACCGCTTGTGACGAGGTGAAGAGACAGGTGGCGGAGGTGCGCCCAA	21641
QY	588	GCAG 591	
Db	21642	GCTG 21645	

RESULT 8
US-09-870-759-129
; Sequence 129, Application US/09870755
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
;; FILE REFERENCE: 870759
;; CURRENT APPLICATION NUMBER: US/09/870,759
;; CURRENT FILING DATE: 2002-01-14
;; PRIOR APPLICATION NUMBER: US 60/208,128
;; PRIOR FILING DATE: 2000-05-30
;; NUMBER OF SEQ ID NOS: 166
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 129
;; LENGTH: 1156
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (61)..(1014)
;; OTHER INFORMATION:
US-09-870-759-129

Query Match 3.5%; Score 50.4; DB 9; Length 1156;
Best Local Similarity 44.0%; Pred. No. 0.012;
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCTCTGCTGACAGAGTGTCTCTGTGATGCTTACCGCACCACGACGAAAGTGGC 167
DB 360 GACCCCGGTGGCGAGAGACGCGGCGACGCTGTCCAGAGAGCTGACAGCGCGCGACAGGC 419
QY 168 AGGTACAGGGCTCTGACGCTGCTTGGCAAGAGGGGGGAGCCCGGACGTCGCA 227
DB 420 CCGGCTGGGGCGGACATGAGAGAGTGTGGCGCGCTGTGACATACCGCGCGAGGT 479
QY 228 GATGCTGAGATCCACCGACGACCCGACGCTGATGTCGAGCTGCGATTTGCGGGCG 287
DB 480 GCGGCGCATGCTGGGCGAGACCGAGAGCTGGGGTGGCTGCTCCACCTGCG 539
QY 288 GCGAGCCTGTGGCGCTTCTTCCGCGCTTACCGGAGGGGCGCTGGCGCGCTGCA 347
DB 540 CAGCTGCGTGAAGCGGCTCTTCCGCGATGCGGATGACTCAGAGAGCCTGGCAGTGA 599
QY 348 GAGAGGCTTGGCGGCGCGCTGCGCGACGCTGCGCGCTGCAATGAGAGCTGGCGGC 407
DB 600 CCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCTGCGCGCATCCGCGCGCGCTGGG 659
QY 408 CCGGCGCGGCGCGCTGAGCGCTTGTCTGCGGAGACGAGAGCGCTTTGATGTCATCT 467
DB 660 GCCCTGTGTGAACAGGGCGCGGTGGGGCGCGCAGCTGTGGGCTCCTTGGCGCGCGCC 719
QY 468 AGCCGACGACCGGACCGGCTCGGGATGAAGAACTGGCTGAGCTGAGATGCGCTGGC 527
DB 720 GCTACAGAGAGCGGGCGCGGCTTGGGGCGAGCGGCTGCGCGCGGATGAGAGATGGG 779
QY 528 AAATCTGAAGTGGCGCTCGGGGGCGCGGGGTGGGACGGGAGGTGCTTCCGCCCTT 587
DB 780 CAGCGGAGCCCGGACCGCTGAGCAGGTGAAGAGCAGAGTGGCGGAGGTGCGGCCAA 839
QY 588 GCAG 591
DB 840 GCTG 843

RESULT 9
US-09-827-854-7
;; Sequence 7, Application US/09827854
;; Patent No. US20020123093A1
;; GENERAL INFORMATION:
;; APPLICANT: Zannis, Vasile
;; APPLICANT: Kyriacos, Kyriacos E.
;; TITLE OF INVENTION: Compounds and methods for lowering
;; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
;; FILE REFERENCE: 07180/004003
;; CURRENT APPLICATION NUMBER: US/09/827,854
;; CURRENT FILING DATE: 2001-04-05
;; PRIOR APPLICATION NUMBER: US 09/679,088
;; PRIOR FILING DATE: 2000-10-04

;; PRIOR APPLICATION NUMBER: US 09/544,386
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 7
;; LENGTH: 1156
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-827-854-7

Query Match 3.5%; Score 50.4; DB 10; Length 1156;
Best Local Similarity 44.0%; Pred. No. 0.012;
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCTCTGCTGACAGAGTGTCTCTGTGATGCTTACCGCACCACGACGAAAGTGGC 167
DB 360 GACCCCGGTGGCGAGAGACGCGGCGACGCTGTCCAGAGAGCTGACAGCGCGCGACAGGC 419
QY 168 AGGTACAGGGCTCTGACGCTGCTTGGCAAGAGGGGGGAGCCCGGACGTCGCA 227
DB 420 CCGGCTGGGGCGGACATGAGAGAGTGTGGCGCGCTGTGACATACCGCGCGAGGT 479
QY 228 GATGCTGAGATCCACCGACGACCCGACGCTGATGTCGAGCTGCGATTTGCGGGCG 287
DB 480 GCGGCGCATGCTGGGCGAGACCGAGAGCTGCGGGTGGCTGCTCCACCTGCG 539
QY 288 GCGAGCCTGTGGCGCTTCTTCCGCGCTTACCGGAGGGGCGCTGGCGCGCTGCA 347
DB 540 CAGCTGCGTGAAGCGGCTCTTCCGCGATGCGGATGACTCAGAGAGCCTGGCAGTGA 599
QY 348 GAGAGGCTTGGCGGCGCGCTGCGCGACGCTGCGCGCTGCAATGAGAGCTGGCGGC 407
DB 600 CCGGCGCGGCGCGCGGCGCGCGCGCGCGCGCGCTGCGCGCATCCGCGCGCGCTGGG 659
QY 408 CCGGCGCGGCGCGCTGAGCGCTTGTCTGCGGAGACGAGAGCGCTTTGATGTCATCT 467
DB 660 GCCCTGTGTGAACAGGGCGCGGTGGGGCGCGCAGCTGTGGGCTCCTTGGCGCGCGCC 719
QY 468 AGCCGACGACCGGACCGGCTCGGGATGAAGAACTGGCTGAGCTGAGATGCGCTGGC 527
DB 720 GCTACAGAGAGCGGGCGCGGCTTGGGGCGAGCGGCTGCGCGCGGATGAGAGATGGG 779
QY 528 AAATCTGAAGTGGCGCTCGGGGGCGCGGGGTGGGACGGGAGGTGCTTCCGCCCTT 587
DB 780 CAGCGGAGCCCGGACCGCTGAGCAGGTGAAGAGCAGAGTGGCGGAGGTGCGGCCAA 839
QY 588 GCAG 591
DB 840 GCTG 843

RESULT 10
US-09-827-854-8
;; Sequence 8, Application US/09827854
;; Patent No. US20020123093A1
;; GENERAL INFORMATION:
;; APPLICANT: Zannis, Vasile
;; APPLICANT: Kyriacos, Kyriacos E.
;; TITLE OF INVENTION: Compounds and methods for lowering
;; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
;; FILE REFERENCE: 07180/004003
;; CURRENT APPLICATION NUMBER: US/09/827,854
;; CURRENT FILING DATE: 2001-04-05
;; PRIOR APPLICATION NUMBER: US 09/679,088
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 09/544,386
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 8
;; LENGTH: 1156
;; TYPE: DNA
;; ORGANISM: Homo sapiens

US-09-827-854-8

Query Match 3.5%; Score 50.4; DB 10; Length 1156;
Best Local Similarity 44.0%; Pred. No. 0.012;
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCTCGCTGACAAAGTGTCTCTGTGGATGCTTACAGGAGCTGCAAGGCGGCGCAGGC 167
DB 360 GACCCCGGTGGCGGAGAGACGCGGCGACGCGCTGTCCAAGAGCTGCAAGGCGGCGCAGGC 419
QY 168 AGTGTACAGGCGCTGTGCAAGGCTGTGCGACAGAGCGGCGGAGACCGCGACGCTGTGCA 227
DB 420 CCGGTGGGCGCGGACATGAGAGAGAGTGTGCGCGCGCTGTGTGCAATGCTGTGCGAGGT 479
QY 228 GATGCTGAAGATTCACCGCAGCGACCGCGACCTGTGACGTGCGCATTTCTGCGGCGC 287
DB 480 GCAGGCCATGTCTCGGCGCAGAGACCGAGAGAGTGTGCGGCGCTGTGCTCCACATCTGG 539
QY 288 GAGCCCTGTGCGCGCTTCTCTCGCGCTTACCGCGAGGGGCGCTGCGCGCGCGCTGCA 347
DB 540 CAAGCTGTGAAGCGGCTCTCTCGCGATGCGCATGACCTGTGCAAGAGCGCGCTGTGCTGA 599
QY 348 GAGAGCGCTGTGCGCGCGCGCTGTGCGCGACCTGTGCGCGCGCTGTGCAATGCGCGC 407
DB 600 CCGGCGCGGCG 659
QY 408 CGGCGCGCGCGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 467
DB 660 GCGCTGTGTGAACAGAGCG 719
QY 468 AGCCAGCAGCG 527
DB 720 GCTACAGAGCG 779
QY 528 AAATGTGAAGTGTGCT 587
DB 780 CAGCGCGAGCA 839
QY 588 GCAG 591
DB 840 GCTG 843

RESULT 11
US-09-827-854-10
; Sequence 10, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kypreos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-10

Query Match 3.5%; Score 50.4; DB 10; Length 1156;
Best Local Similarity 44.0%; Pred. No. 0.012;
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCTCGCTGACAAAGTGTCTCTGTGGATGCTTACAGGAGCTGCAAGGCGGCGCAGGC 167
DB 360 GACCCCGGTGGCGGAGAGACGCGGCGACGCGCTGTCCAAGAGCTGCAAGGCGGCGCAGGC 419
QY 168 AGTGTACAGGCGCTGTGCAAGGCTGTGCGACAGAGCGGCGGAGACCGCGACGCTGTGCA 227
DB 420 CCGGTGGGCGCGGACATGAGAGAGAGTGTGCGCGCGCTGTGTGCAATGCTGTGCGAGGT 479
QY 228 GATGCTGAAGATTCACCGCAGCGACCGCGACCTGTGACGTGCGCATTTCTGCGGCGC 287
DB 480 GCAGGCCATGTCTCGGCGCAGAGACCGAGAGAGTGTGCGGCGCTGTGCTCCACATCTGG 539
QY 288 GAGCCCTGTGCGCGCTTCTCTCGCGCTTACCGCGAGGGGCGCTGCGCGCGCGCTGCA 347
DB 540 CAAGCTGTGAAGCGGCTCTCTCGCGATGCGCATGACCTGTGCAAGAGCGCGCTGTGCTGA 599
QY 348 GAGAGCGCTGTGCGCGCGCGCTGTGCGCGACCTGTGCGCGCGCTGTGCAATGCGCGC 407
DB 600 CCGGCGCGGCG 659
QY 408 CGGCGCGAGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGATCT 467
DB 660 GCGCTGTGTGAACAGAGCG 719
QY 468 AGCCAGCAGCG 527
DB 720 GCTACAGAGCG 779
QY 528 AAATGTGAAGTGTGCT 587
DB 780 CAGCGCGAGCA 839
QY 588 GCAG 591
DB 840 GCTG 843

DB 360 GACCCCGGTGGCGGAGAGACGCGGCGACGCGCTGTCCAAGAGCTGCAAGGCGGCGCAGGC 419
QY 168 AGTGTACAGGCGCTGTGCAAGGCTGTGCGACAGAGCGGCGGAGACCGCGACGCTGTGCA 227
DB 420 CCGGTGGGCGCGGACATGAGAGAGAGTGTGCGCGCGCTGTGTGCAATGCTGTGCGAGGT 479
QY 228 GATGCTGAAGATTCACCGCAGCGACCGCGACCTGTGACGTGCGCATTTCTGCGGCGC 287
DB 480 GCAGGCCATGTCTCGGCGCAGAGACCGAGAGAGTGTGCGGCGCTGTGCTCCACATCTGG 539
QY 288 GAGCCCTGTGCGCGCTTCTCTCGCGCTTACCGCGAGGGGCGCTGCGCGCGCGCTGCA 347
DB 540 CAAGCTGTGAAGCGGCTCTCTCGCGATGCGCATGACCTGTGCAAGAGCTGTGCGAGTGA 599
QY 348 GAGAGCGCTGTGCGCGCGCGCTGTGCGCGACCTGTGCGCGCGCTGTGCAATGCGCGC 407
DB 600 CCGGCGCGGCG 659
QY 408 CGGCGCGAGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGATCT 467
DB 660 GCGCTGTGTGAACAGAGCG 719
QY 468 AGCCAGCAGCG 527
DB 720 GCTACAGAGCG 779
QY 528 AAATGTGAAGTGTGCT 587
DB 780 CAGCGCGAGCA 839
QY 588 GCAG 591
DB 840 GCTG 843

RESULT 12
US-09-827-854-12
; Sequence 12, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kypreos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-12

Query Match 3.5%; Score 50.4; DB 10; Length 1156;
Best Local Similarity 44.0%; Pred. No. 0.012;
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCTCGCTGACAAAGTGTCTCTGTGGATGCTTACAGGAGCTGCAAGGCGGCGCAGGC 167
DB 360 GACCCCGGTGGCGGAGAGACGCGGCGACGCGCTGTCCAAGAGCTGCAAGGCGGCGCAGGC 419
QY 168 AGTGTACAGGCGCTGTGCAAGGCTGTGCGACAGAGCGGCGGAGACCGCGACGCTGTGCA 227
DB 420 CCGGTGGGCGCGGACATGAGAGAGAGTGTGCGCGCGCTGTGTGCAATGCTGTGCGAGGT 479
QY 228 GATGCTGAAGATTCACCGCAGCGACCGCGACCTGTGACGTGCGCATTTCTGCGGCGC 287
DB 480 GCAGGCCATGTCTCGGCGCAGAGACCGAGAGAGTGTGCGGCGCTGTGCTCCACATCTGG 539
QY 288 GAGCCCTGTGCGCGCTTCTCTCGCGCTTACCGCGAGGGGCGCTGCGCGCGCGCTGCA 347
DB 540 CAAGCTGTGAAGCGGCTCTCTCGCGATGCGCATGACCTGTGCAAGAGCTGTGCGAGTGA 599
QY 348 GAGAGCGCTGTGCGCGCGCGCTGTGCGCGACCTGTGCGCGCGCTGTGCAATGCGCGC 407
DB 600 CCGGCGCGGCG 659
QY 408 CGGCGCGAGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGATCT 467
DB 660 GCGCTGTGTGAACAGAGCG 719
QY 468 AGCCAGCAGCG 527
DB 720 GCTACAGAGCG 779
QY 528 AAATGTGAAGTGTGCT 587
DB 780 CAGCGCGAGCA 839
QY 588 GCAG 591
DB 840 GCTG 843

```

Db 480 GCAGCCCATCTGCGCCAGACACCGAGCTGCGGGTGGCCTTCCTCCACTGCG 539
QY 288 GCAGCCCTGTGCGCGCTTCTCCGCGCTACCGGAGGGGCGCTGGCGCGCGCTGCA 347
Db 540 CAGCTGCTGAGGCGGCTCTCCGCGATGCCGATGACTGCAAGAGGCCCTGGGAGTGA 599
QY 348 GAGGAGCTTGGCGGCGCGCTCGCCGACACTCGATGCCCTGCACTGAGCTGCGCGC 407
Db 600 CAGGCGCGGGGCGCGCGCGCGCGCGCGCGCTGAGCGCCGCTGAGCGCCGAGCGCG 659
QY 408 CGGCGCGGCGCGCTGAGCGCTTCTGCTGGCGGAGAGAGCGCTGTTTGAATTGCT 467
Db 660 GCCCTGTGTGAACAGGAGCGCGCTGGCGCGCGCCGCTGCTGTGGCTTCCCTGGCGCG 719
QY 468 AGCCGACGAGCGCGCGCGCTCGGAGATGAAGACTGAGCTGAGGATGGCGCTGCG 527
Db 720 GCTACAGAGAGCGCGCGCGCGCTGAGGCGAGCGCGCTGGCGCGCGAGTGGAGTGG 779
QY 528 AAATCTGAAGTGGCGCTTCGCGGCGCGCGCGCGCGCGCGCGCGCTTCCGCGCGCT 587
Db 780 CAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 839
QY 588 GCAG 591
Db 840 GCTG 843

```

RESULT 13

```

US-10-044-090-454
; Sequence 454, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 454
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 2514988CB1
; LOCATION: 46
; OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-454

```

```

Query Match 3.5%; Score 50.4; DB 12; Length 1291;
Best Local Similarity 44.0%; Pred. No. 0.013;
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

```

```

QY 108 GTCTGCTGAGCAAGGAGTGTCTGTGAGATGCTTACGCGACCCCGACGAGAGTGGC 167
Db 485 GACCCCGGTGGGAGGAGAGCGCGGCACGGCTTCAAGAGCTGAGGCGGCGAGCG 544
QY 168 AGTGTACAGGCTCTGAGAGCTGCTTGGCAGAGAGCGCGGAGCCCGGAGCTGTGCA 227
Db 545 CCGGCTGGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGAGTACCGCGCGAGGT 604
QY 228 GATCTGAAGATTCACGCGAGCGACCGCGAGCTGATGTGAGCTGCGATTTCGCGGCG 287
Db 605 GCAGGCGATGCTCGCGCAGACCGAGAGCTGCGGCTGCGCTTCCCTCCACTGCG 664
QY 288 GCAGCGCTGAGCGCGCTTCTCGCGGCTACCGGAGGCGCGCTGGCGCGCGCTGCA 347
Db 665 CAGCTGCTGAGGCGCTTCTCGCGGATGCCGATGACTGCAAGAGCGCTGGGAGTGA 724
QY 348 GAGAGCGCTGGCGCGCGCTGCGCGCACTCGTGCCTGCACTGAGAGCTGCGCGC 407

```

```

Db 725 CAGGCGCGGCGCGCGAGGCGCGGAGCGCGCTTACGCGCATCCGAGCGCGCTGG 784
QY 408 CGGCGCGGAGCGGCTGAGCGCTTGTGCGCGGAGAGAGCGCTGTTTGAATTGCACT 467
Db 785 GCCCTGTGTGAACAGGAGCGCGCTGGCGCGCGCGCACTGTGGCTTCCCTGGCGCG 844
QY 468 AGCCGACGAGCGCGCGCTCGGAGATGAAGACTGAGCTGAGAGATGCGCTGCG 527
Db 845 GCTACAGAGAGCGCGCGCGCGCGCGCGCGCTGAGCGCGCGCGCGCGAGTGGAGTGG 904
QY 528 AAATCTGAAGTGGCGCTTCGCGGCGCGCGGAGTGGCGAGCGGAGAGTGTGCGCGCT 587
Db 905 CAGCGGAGCGCGGAGCGCGCTGAGCAGAGTGAAGAGAGAGTGGCGGAGGTGCGCGCAA 964
QY 588 GCAG 591
Db 965 GCTG 968

```

RESULT 14

```

US-09-772-304A-1
; Sequence 1, Application US/09772304A
; Patent No. US20020146785A1
; GENERAL INFORMATION:
; APPLICANT: Maheshi, L.H.
; APPLICANT: Tripathi, G.
; APPLICANT: Ramchandrar, T.V.N.
; APPLICANT: Rawal, S.K.
; TITLE OF INVENTION: NOVEL ESCHERICHIA COLI HAVING ACCESSION
; FILE REFERENCE: A33943 066123.0103
; CURRENT APPLICATION NUMBER: US/09/772,304A
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4826
; TYPE: DNA
; ORGANISM: Streptomyces aureofaciens
US-09-772-304A-1

```

```

Query Match 3.5%; Score 50.4; DB 10; Length 4826;
Best Local Similarity 42.9%; Pred. No. 0.018;
Matches 368; Conservative 0; Mismatches 481; Indels 9; Gaps 2;

```

```

QY 103 GTGAGTCTCTGCTGAGCAAGTGTCTGTGAGTCTTACGCGACCCCGACGAGAG 162
Db 1377 GTGCGCGGCGCGGTGAGCGCGCGGCTGTGTGCTGCGCGCTCGAGGCGCGCTGTACGCG 1436
QY 163 GTGCGAGTGTACAGGCGCTGTGAGGCTGCTGCGCGCGCGAGAGCGCGCGAGCTG 222
Db 1437 GTGTGTGCGCGCGCTGTGAGGCGCTGTGTGCGCGCGCGCGCTGTGCGCGCGCGCG 1496
QY 223 CTCAGATGTGAAGATTCACCGCAGAGCGCGCGAGCTGAGCTGCGATTCG 282
Db 1497 CCGCGTGGGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1556
QY 283 GGGCGGAGCGCTGTGCGCGCTTCTTCCGCGCTTACCGCGAGGCGCGCGCGCGCGCG 342
Db 1557 GCGCGGCTTGTGGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1616
QY 343 CTGCAAGAGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 402
Db 1617 CTGACGAGTGTGTCTGAGGAGCAGACTGCGCGCGGAGACCGGATGCGCTTGGCGCG 1676
QY 403 CGGCGCGCGCGCGAGCGCTGAGCGCTTGTGCTGCGCGAGAGAGCGCTGTGAGTTCG 462
Db 1677 ACGGAGCGCTGAGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1736
QY 463 ATCTTACCGCAGAGCGCGCGCGCTTCCGAGTGAAGAACTGAGCTGAGAGATGCG 522
Db 1737 GCCTGTGTGCTGCGCGCGCTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1796

```

Qy 523 CTGCGAAATCTGAAGTCGCGCTCGGGGGCCGAGGAGTGCCTTCGACC 582
Db 1797 CGGCGCGAGGGGGCTACTGCGCGCGACGCTG-----TGTGCTGGCCACTGGGG 1850
Qy 583 CCCTTGACGCCCCCGGTCCCTCTCTGTGAGAGTGAACCGCCCGCCGCTCACTT 642
Db 1851 CGGCTGGGGGCTCCGAGATGTGCTTCCCGGAGGCGTACACGCTCCGATGACGGTGGCC 1910
Qy 643 GCCAGACTTTTCTGTTCCAGGGTCAAGCCTGTAAGTAATCGGCGCTGAGCTTCAAGAGAC 702
Db 1911 GCGCTGCTGTGTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1970
Qy 703 CACAGACGTTTCGCGCGCTGTGTGAGTCTCAATGACGCGAGAGTGGGCGCTCACTGACG 762
Db 1971 TACGGGCGG---GGGCTGGCGCGCGACGCTGCTGCCAGCTTGGCGGTGCGCTGACCGAC 2027
Qy 763 CGAGCTGCGGGGCGCTGCGGAGCCCGCGCTGACTGCTGCTGCTTACAGATGACAGCGC 822
Db 2028 CCGGACTGCTCAGGCGCTGTGCTGTGGGAGCGCGCGCTGTGATCACCTGCTCGGC 2087
Qy 823 GAGGACTGTACGAGCAGGCGCTTCAGACTGTCGCGCGCTTGTGACGCGCGAGGCGCGC 882
Db 2088 GCGCGCCACCGGCTCCAGGCGCTGCTGCTGCTGCGCGGAGCGTGTGCTGCTGCTGCGC 2147
Qy 883 CGCGCCAGCTGCAAGCGCTGTGTGAGGAGCACTCGAGAGAACGAGCTACACGCTGCGCA 942
Db 2148 CTGCAACGAGCTGCGCCCGTACGCTGTGAGAGTGGGGGTGGCGCTCCCGCTGCTCCCG 2207
Qy 943 GAGACTGTGCTGGCGCTG 960
Db 2208 CCGCGCTGCGCGCGCTG 2225

RESULT 15

US-09-976-059-1
; Sequence 1, Application US/09976059
; Patent No. US20020164747A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Stafia, Alfredo
; TITLE OF INVENTION: Genes and Proteins for Biosynthesis of Ramoplanin
; FILE REFERENCE: 3019-PC1
; CURRENT APPLICATION NUMBER: US/09/976,059
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent version 3.0
; SEQ ID NO 1
; LENGTH: 88421
; TYPE: DNA
; ORGANISM: Actinoplanes sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2077)..(3078)
; OTHER INFORMATION: ORF 1; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (3118)..(4032)
; OTHER INFORMATION: ORF 2; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (4038)..(5048)
; OTHER INFORMATION: ORF 3; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (4038)..(5048)
; OTHER INFORMATION: ORF 3; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (6665)..(5814)
; OTHER INFORMATION: ORF 4; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (7703)..(6693)
; OTHER INFORMATION: ORF 5; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (9464)..(8130)
; OTHER INFORMATION: ORF 6; negative strandedness

; NAME/KEY: misc_feature
; LOCATION: (9691)..(10761)
; OTHER INFORMATION: ORF 7; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (12751)..(10829)
; OTHER INFORMATION: ORF 8; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (13617)..(12802)
; OTHER INFORMATION: ORF 9; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (15203)..(13614)
; OTHER INFORMATION: ORF 10; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (15591)..(15863)
; OTHER INFORMATION: ORF 11; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (15860)..(19035)
; OTHER INFORMATION: ORF 12; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (19032)..(39713)
; OTHER INFORMATION: ORF 13; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (39713)..(65800)
; OTHER INFORMATION: ORF 14; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (65826)..(66530)
; OTHER INFORMATION: ORF 15; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (66546)..(67370)
; OTHER INFORMATION: ORF 16; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (67384)..(70059)
; OTHER INFORMATION: ORF 17; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (70059)..(70662)
; OTHER INFORMATION: ORF 18; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (70659)..(71906)
; OTHER INFORMATION: ORF 19; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (73439)..(71964)
; OTHER INFORMATION: ORF 20; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (74216)..(73563)
; OTHER INFORMATION: ORF 21; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (75424)..(74213)
; OTHER INFORMATION: ORF 22; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (75535)..(76464)
; OTHER INFORMATION: ORF 23; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (78110)..(76449)
; OTHER INFORMATION: ORF 24; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (79864)..(78107)
; OTHER INFORMATION: ORF 25; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (81624)..(79861)
; OTHER INFORMATION: ORF 26; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (81909)..(81682)
; OTHER INFORMATION: ORF 27; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (82346)..(82062)
; OTHER INFORMATION: ORF 28; negative strandedness
; NAME/KEY: misc_feature
; LOCATION: (82587)..(84446)
; OTHER INFORMATION: ORF 29; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (84481)..(85548)
; OTHER INFORMATION: ORF 30; positive strandedness
; NAME/KEY: misc_feature

; LOCATION: (85556)..(86845)
; OTHER INFORMATION: ORF 31; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (87372)..(86803)
; OTHER INFORMATION: ORF 32; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (87494)..(88420)
; OTHER INFORMATION: ORF 33; positive strandedness; N-terminus only
US-09-976-059-1

Query Match 3.5%; Score 50.2; DB 9; Length 88421;
Best Local Similarity 46.4%; Pred. No. 0.041;
Matches 163; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 97 CTGTTTGAGAGTCTCTGACAGAGTGTCTCTGCGATGCTTACGGCACCCCCAG 156
DB 65955 CTGCTCGACGACCTTACGCGGACCTGCTCGCGGCGCGGACCTGCGCTTACCTTC 66014
QY 157 CAGAGGTGGCAGTGTACAGGCTCTGACAGCTGCTTGGCAGAGAGCGGAGGCCG 216
DB 66015 GCCGACGAGGTGGGCGGGGGGCGGGGTGGCGCTTTCGGGACACGCTCGGGCCGTC 66074
QY 217 GAGTGTCTGACAGTGTGAAGATCACCGACGACCCGACGCTGATCTGCACTGCGA 276
DB 66075 CTGCGGTACGAGATGCGCACGCGCTGTGTGCGCACGCGAGGTGGCTGTGAGCGGCTC 66134
QY 277 TTCTGCGGGGCGGACCGCTGTGGCGCTTCCGCGCTTACCGGCGAGGGGGGCGTGGC 336
DB 66135 TTGCTCAGCGGCTGCGCCGATCCCTTGACCCCTTGACACCGGCGAGGGGCTTGAC 66194
QY 337 GCCGCGCTGACAGAGACCTGCGGCGCGCTGCGCCAGCACTGCTGCGCGTGGCACTG 396
DB 66195 GACGAGAGATTCTGCTGCGGGTGGCGAGTTGCGCGTTACGACCAAGAGGCGCTGCGC 66254
QY 397 GAGTGTGCGCGCGCGCGCGCGAGCGCTTGTCTGCGGCGAGCGAG 447
DB 66255 GATCCGACATGCGCGAGCTGATCTGCGCGCTGCGCGCGACGTCGAG 66305

Search completed: February 3, 2003, 16:17:35
Job time: 194 secs

THIS PAGE BLANK (USPTO)